

Estimación de coordenadas WGS-84 de orillas mediante procesamiento digital de mediciones obtenidas con radar

JHONATHAN A. POSADA

**Universidad del Norte
División de Ingenierías
Barranquilla, noviembre de 2010**

Estimación de coordenadas WGS-84 de orillas mediante procesamiento digital de mediciones obtenidas con radar

JHONATHAN ALBERTO POSADA CÁCERES
Ingeniero Electrónico

Trabajo de grado presentado como requisito para optar al título de
Magíster en Ingeniería Electrónica

Director
Ph.D JUAN CARLOS VÉLEZ DÍAZ

UNIVERSIDAD DEL NORTE
DIVISIÓN DE INGENIERÍA
PROGRAMA DE MAESTRÍA EN INGENIERÍA ELECTRÓNICA
BARRANQUILLA
2010

Aprobado por el profesorado de
la División de Ingenierías en
cumplimiento de los requisitos
exigidos para otorgar el título de

**MAGISTER EN INGENIERÍA
ELECTRÓNICA**

**Ph.D JUAN CARLOS VÉLEZ DÍAZ
DIRECTOR**

CORRECTOR

CORRECTOR

**PH.D CRISTIAN QUINTERO
COORDINADOR DE PROGRAMA**

Barranquilla, 18 de noviembre de 2010

AGRADECIMIENTOS

A Dios, por las oportunidades que se presentaron para desarrollar esta tesis como lo esperábamos.

A mi madre y mi padre, quienes me han apoyado en este tiempo con atenciones, recursos y esfuerzos también de su parte.

A Juan Carlos Vélez, doy mis agradecimientos especiales por su confianza en mí, no sólo al confiarme la idea inicial de este proyecto, sino por el apoyo al incluirme en todos sus inventivas. Profe, de veras que gracias, y tengo la certeza de que su enfoque es el adecuado, la investigación debe aportarle algo a nuestro entorno, al País, a la Ciudad, ¿Quién más va a resolver nuestros problemas?

Al Equipo de UNINORTE, lo menciono especialmente pues ha sido el que nos ha acompañado en CORMAGDALENA: el ingeniero Manuel Alvarado, Holbert, Francia y, sin duda, la Ingeniera Consuelo Castilla que nos ha apoyado en el desarrollo del SNS.

A Tania, a quien dejé de último en estos agradecimientos como prueba de que el orden de mención no es discriminación personal; a ella, la que me ha acompañado en los momentos felices y duros del proyecto, mi Tanya, quien le ha dado color a mis días, gracias por tu apoyo en estos meses, por tu comprensión y por tu muy valiosa asesoría y aportes en el proyecto.

TABLA DE CONTENIDO

pág.

ORGANIZACIÓN DEL DOCUMENTO.....	1
1 INTRODUCCIÓN	2
1.1 PROBLEMÁTICA Y ENFOQUE DE ESTA TESIS	3
1.1.1 Ámbito Socio Económico.....	3
1.1.2 Ámbito Científico Técnico	5
1.2 OBJETIVOS DE LA TESIS	6
1.2.1 Objetivo General	6
1.2.2 Objetivos Específicos	6
1.3 APORTES ESPERADOS	7
1.3.1 A nivel Científico - Tecnológico.....	7
1.3.2 A nivel de Productividad y Competitividad.....	7
1.3.3 En el medio Ambiente y la Sociedad.....	8
2 MARCO REFERENCIAL DE LA INVESTIGACIÓN	9
2.1 MARCO TEÓRICO.....	9
2.1.1 Acerca de la Georreferenciación.....	9
2.1.2 Acerca del Sensado Remoto	10
2.2 MARCO CONCEPTUAL	11
2.2.1 Representación de Sistemas de Información Geográfica (SIG).....	11
2.2.2 Sistemas de Referencia Geográfica.....	12
2.2.3 Estimación Bayesiana de Parámetros	14
2.2.4 GPS: Nociones Básicas de Operación.....	17
2.2.5 Sensor de Azimut Electrónico: Nociones Básicas de Operación	19
2.2.6 Radar: Nociones Básicas de Operación.....	20
2.2.7 Detección de Líneas y Bordes en Imágenes.....	21
2.3 ESTADO DE DESARROLLO DEL TEMA	22
2.3.1 Trabajos y Técnicas Actuales en Georreferenciación y Sensado Remoto Aplicado	22
2.3.2 Trabajos y Técnicas Actuales en Estimación de Parámetros y Variables	25
3 IMPLEMENTACIÓN DEL SISTEMA DE GEORREFERENCIACIÓN	28
3.1.1 Dispositivo de Enlace al Radar	29
3.1.2 Software de Integración y Adquisición Sincronizada de Datos	33
3.1.3 Algoritmo de Georreferenciación	37
4 DESCRIPCIÓN TEÓRICA DEL SISTEMA PROPUESTO	38
4.1 MODELO MATEMÁTICO PARA LAS COORDENADAS DE SALIDA.....	38
4.2 ESTIMACIÓN DE COORDENADAS DE LA ORILLA	43
4.2.1 Generación de Subgrupos (Cluster's) para Puntos Significativos de la Orilla.....	45

4.2.2	Reducción de Subgrupos a Coordenadas de menor Incertidumbre	48
4.3	MÉTRICAS DE DESEMPEÑO PARA EL SISTEMA	52
4.3.1	Cálculo de la Resolución	52
4.3.2	Cálculo de estadísticos del error	53
5	RESULTADOS OBTENIDOS Y ANÁLISIS.....	55
5.1	RESULTADOS PRELIMINARES DEL SISTEMA SIN ESTIMACIÓN.....	56
5.1.1	Calibración de Mediciones del Radar.....	56
5.1.2	Georreferenciación Inicial con Radar	57
5.2	OPTIMIZACIÓN DE LA CANTIDAD DE SUBGRUPOS GENERADOS	58
5.3	SELECCIÓN DEL ESTIMADOR PARA REDUCCIÓN DE INCERTIDUMBRE	62
5.4	DESEMPEÑO FRENTE A OTRAS TÉCNICAS DE GEORREFENCIACIÓN	64
6	CONCLUSIONES	68
	RECOMENDACIONES.....	70
	REFERENCIAS Y BIBLIOGRAFÍA	71
	ANEXOS	76
1.	ALGORITMO DE GEORREFERENCIACIÓN INVERSA PARA UBICAR COORDENADAS RUIDOSAS DE UN MISMO PUNTO	77
2.	LISTA DE SIMBOLOS UTILIZADOS.....	79

ORGANIZACIÓN DEL DOCUMENTO

La obra ha sido organizada en seis capítulos principales cuyos contenidos son:

- **Capítulo 1**, se presenta la introducción de la investigación relatando el problema a solucionar tanto a nivel científico como socioeconómico. Se establecen objetivos, alcances y aportes esperados.
- **Capítulo 2**, contiene el marco referencial de la investigación, donde se inicia estableciendo y delimitando las áreas de conocimiento involucradas, se presentan los conceptos necesarios para la comprensión de la tesis y, finalmente, se relata el estado de desarrollo mediante los trabajos y técnicas relacionados a la aplicación.
- **Capítulo 3**, expone la implementación práctica y computacional utilizada para evaluar el funcionamiento de la propuesta, explicando las interconexiones físicas y virtuales, así como una descripción de las soluciones en hardware usadas.
- **Capítulo 4**, presenta el desarrollo teórico y matemático de la solución propuesta a través de los modelos y razonamientos que lo conforman.
- **Capítulo 5**, detalla las pruebas realizadas, los resultados obtenidos y el correspondiente análisis.
- **Capítulo 6**, expone las conclusiones a partir de la revisión de los resultados obtenidos contra los objetivos planteados y las limitaciones de la implementación.

Finalmente, se encontrarán las recomendaciones para continuar la mejora de este trabajo, las referencias bibliográficas y anexos.

1 INTRODUCCIÓN

La navegación se constituye como una de las actividades humanas de mayor importancia, vinculada especialmente con el transporte de personas y mercancías. En la actualidad, dado el escalamiento al que se han llevado tanto las embarcaciones como las cargas transportadas, se busca garantizar la seguridad con el apoyo de la tecnología. Históricamente se utilizaban las cartas de navegación en papel para ubicar la ruta de una embarcación en un contexto geográfico. Los elementos de apoyo han evolucionado con el transcurso del tiempo tendiendo hacia la navegación electrónica con recursos para la asistencia, control y monitoreo de la embarcación y sistemas de navegación satelital basados en cartas náuticas digitales.

En general, la realización de cartas para la navegación es costosa por la gran cantidad de levantamientos cartográficos y batimétricos requeridos, aunque para la navegación marítima no representa un problema debido a la larga vigencia de estas y la gran profundidad del mar en la mayor parte de su extensión. No ocurre de la misma manera en el caso de las cartas náuticas fluviales, en las que el levantamiento cartográfico de orillas e islas de ríos se constituye en una actividad importante y frecuente; En el caso fluvial, los procesos de sedimentación y erosión provocan grandes variaciones en la ubicación de orillas, islas y el canal navegable, de ahí que se presenten altos costos para el mantenimiento de cartas náuticas actualizadas.

En consecuencia, existe la necesidad de implementar nuevas alternativas técnicas para el levantamiento cartográfico de orillas e islas con mayor agilidad a las existentes y costos de operación que permitan una utilización intensiva y continua. La presente investigación desarrolla los aspectos prácticos y teóricos necesarios para la utilización de radares comerciales y sistemas de posicionamiento global (GPS) con miras a una técnica de georreferenciación remota de orillas.

Otra aplicación adicional a esta propuesta, se encontrará en el área de la hidrografía, donde se requiere un seguimiento detallado a la evolución de la erosión costera. Aquí, la información de cartas náuticas y mapas oficiales no es siquiera tenida en cuenta, pues se hace necesario el seguimiento de la línea de costa con técnicas que brinden resoluciones en el orden del metro, con cubrimiento de largas extensiones. La precisión y exactitud que se obtenga del presente desarrollo, determinará la viabilidad de extender su uso para el levantamiento topográfico de alta precisión requerido en el seguimiento de la dinámica costera.

1.1 PROBLEMÁTICA Y ENFOQUE DE ESTA TESIS

1.1.1 Ámbito Socio Económico

Dentro de los planes estratégicos de desarrollo del estado colombiano y la región Caribe, se encuentra la necesidad de mejorar la competitividad de todo su sistema de transporte, con miras a una mejor integración de sus centros de producción con los puertos de exportación y los países vecinos. Se ha reconocido que por las características territoriales y topográficas del país y los bajos costos del transporte fluvial, este medio tiene un alto potencial para el movimiento de pasajeros y carga a gran escala (DNP, 2004). En este sentido, se estableció como meta la consolidación del transporte fluvial en las principales cuencas del país, mediante un incremento del 80% de navegación para el 2019 con respecto al nivel de navegación actual, así como un aumento del 3% al 10% de la carga transportada en el país (DNP, 2004).

La red fluvial del país cuenta con una longitud de 18.225 Km. aptos para navegación menor permanente durante todo el año, de los cuales cuenta con 7.063 Km. (39%) que permiten navegación mayor permanente y 4210 Km. que permiten navegación transitoria (Mintransporte, 2000). Además de las obras civiles correspondientes, la escalabilidad de la navegación en toda la red requiere como factor determinante la prevención de accidentes mediante una adecuada señalización (DNP, 2004). Sin embargo, de los 18.225 Km. de la red fluvial escasamente 550 Km. se encuentran señalizados.

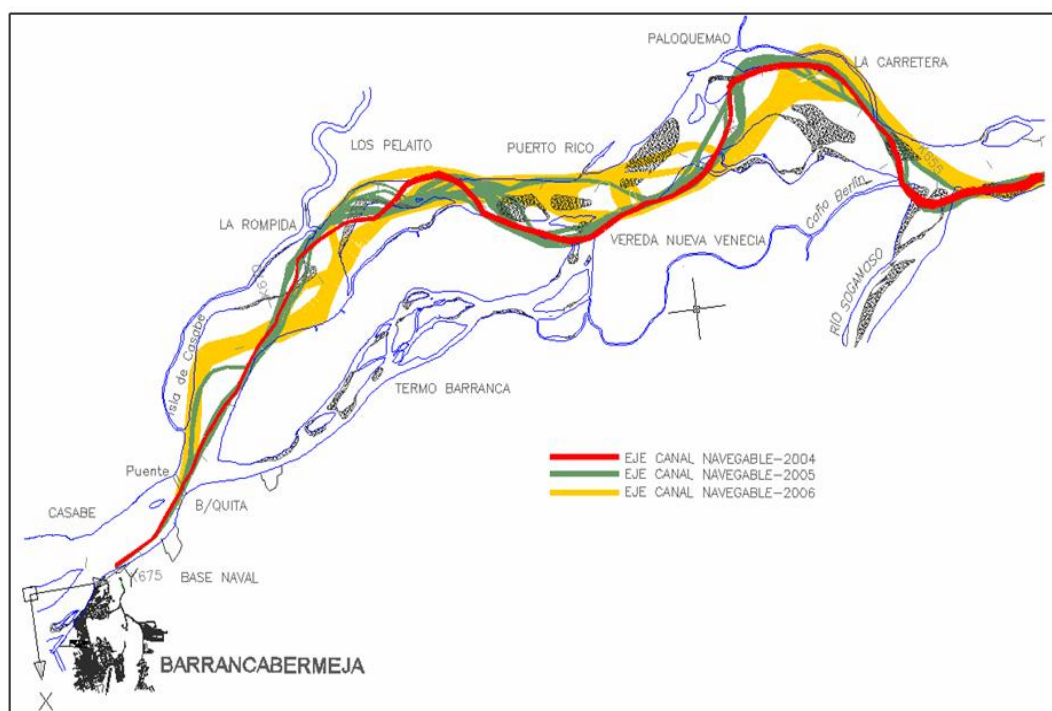
El tramo entre Barranquilla y Puerto Salgar cuenta con señalización satelital realizada por Cormagdalena y la Universidad del Norte mediante el proyecto “Diseño e Implementación del Sistema de Asistencia Satelital a la Navegación en el Río Magdalena entre Puerto Salgar (K921) y Barranquilla (K0) – SNS”, ganador del Premio Nacional de Ingeniería “Lorenzo Codazzi” 2007. En esa realización, se demostró una mejoría en los tiempos de navegación en el orden del 30%, pronosticando la iniciativa como una de las más viables para el apoyo en la recuperación de la confianza y la navegación en las vías hídricas nacionales (Alvarado, 2009).

Sin embargo, la poca actualización de la cartografía de orillas afecta la seguridad, claridad y confianza de los usuarios al utilizar los sistemas de asistencia satelital a la navegación, como lo ha demostrado la investigación del proyecto “Actualización del Sistema de Navegación Satelital” citado en (Posada y Velez, 2010), donde los usuarios ante la responsabilidad de dirigir embarcaciones con cargas mayores a 600 toneladas, se abstienen de reemplazar sus sentidos con la herramienta de posicionamiento satelital. El mantenimiento a costos razonables de una referencia geográfica actualizada de orillas e islas del río con cartas náuticas digitales de mayor calidad representa uno de los retos para la incorporación masiva de las tecnologías de navegación electrónica dentro del transporte fluvial colombiano.

En su estado de desarrollo actual, la georreferenciación de la línea de orilla para ríos, islotes y costas, se realiza mediante técnicas que tienen un alto costo por kilómetro y poca agilidad en su elaboración. Por esta razón, frecuentemente se presenta que la información geográfica de orillas no sea consecuente con la dinámica de éstas, provocando que cartas náuticas, estudios hidrodinámicos del río, entre otros, aparezcan obsoletos.

Por ejemplo, en la construcción de cartas náuticas digitales fluviales en Colombia, dados los altos costos del proceso, la referencia geográfica de orillas e islas se realiza en promedio cada tres años, periodo en el cual el cauce del río ha cambiado tanto que el cauce principal o canal navegable del río se superpone con el trazo de la referencia geográfica de las orillas (véase Figura 1.1). En la Figura citada debe notarse como las actualizaciones del canal navegable en los años 2004, 2005, 2006 se superponen a las líneas azules que representan la ubicación de las orillas según su última georreferenciación.

Figura 1.1. Posición del canal navegable entre 2004-2006 en un sector trezado del río Magdalena.



Tomada de (Alvarado, 2009).

La situación mencionada se presenta especialmente a lo largo de todos los sectores trezados del río Magdalena, que representan el 50% de su longitud y se caracterizan por su dispersión en varios canales y presencia de varias islas con alta movilidad (Alvarado, 2009). Dado los cambios causados por la hidrodinámica en ríos trezados, la vigencia de

la información geográfica en estos sectores es de alrededor de tres semanas, y la ubicación del canal navegable varía en cuestión de días.

Teniendo en cuenta la problemática planteada, las posibilidades de crecimiento en la utilización de la red fluvial nacional y las necesidades de mantenimiento de los sistemas de navegación satelital, se hace indispensable encontrar metodologías para la creación de cartas náuticas fluviales de manera más ágil y económica, con la calidad en resolución y exactitud idóneas para dichas aplicaciones.

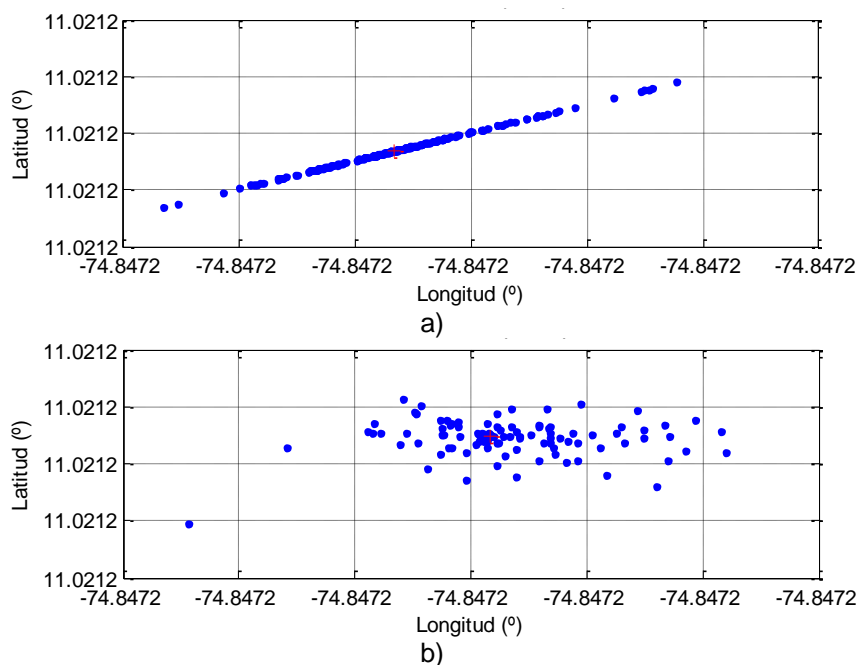
1.1.2 Ámbito Científico Técnico

El proyecto de grado de (Jiménez & Posada, 2007) mostró la posibilidad teórica de abordar el problema de georreferenciación de orillas e islas mediante su sensado remoto con radar de corto alcance desde una embarcación dotada con GPS, transformando geométricamente la información para obtener la ubicación de las orillas. En dicho proyecto se realizó la simulación del proceso de navegación a través de canales navegables y de la información que en teoría debían entregar cada uno de los instrumentos. Si bien los resultados fueron positivos, aún quedó pendiente la implementación real del sistema y la solución del problema de sincronización, redundancia e imprecisión asociado a equipos de medición reales.

En ese sentido, se mantiene la necesidad de estimar la “coordenada geográfica más exacta” a partir de un conjunto finito de datos ruidosos provenientes de un radar, un localizador satelital GPS y una brújula electrónica. Adicionalmente, el patrón de dispersión de las coordenadas de salida depende de cual instrumento está presentando mayores desviaciones en un momento dado. En la Figura 1.2a se presenta el patrón de dispersión de las coordenadas estimadas referentes a un mismo punto geográfico cuando la fuente de mayor dispersión en las coordenadas es la imprecisión del radar. Así mismo, cuando la mayor incertidumbre es ocasionada por la imprecisión del GPS, el patrón tiende a comportarse como el presentado en la Figura 1.2b. El patrón para una brújula con error dominante es similar al de la Figura 1.2a, pero las muestras tienden a distribuirse de manera perpendicular al presentado. Esta situación ocurre en cada una de las coordenadas que conforman a una orilla.

La resolución del problema planteado implica el abordaje de la teoría estadística relacionada con estimadores bayesianos (Vaseghi, 2000), según la cual es posible acercarse al valor real de un parámetro en un sistema si se conoce a priori la distribución de probabilidad de la aleatoriedad que afecta a las mediciones. Por la forma en que se recolecta la información en este sistema, previamente se debe encontrar el particionamiento adecuado del conjunto de coordenadas de salida en subgrupos que hagan referencia a un mismo punto geográfico sobre la orilla. En consecuencia, la presente investigación determinará si es posible calcular con una exactitud aceptable las coordenadas geográficas de orillas fluviales mediante su sensado remoto con radar, la transformación matemática de dicha información y el posterior tratamiento matemático y estadístico de dicha información.

Figura 1.2. a) Patrón de dispersión de coordenadas geográficas con error de radar dominante.
b) Patrón de dispersión de coordenadas geográficas con error del GPS dominante.



1.2 OBJETIVOS DE LA TESIS

1.2.1 Objetivo General

Diseñar e implementar un algoritmo de procesamiento para la referenciación geográfica WGS84 de orillas detectadas con radar.

1.2.2 Objetivos Específicos

- Establecer la relación matemática entre las variables de entrada y las coordenadas WGS84 obtenidas.
- Desarrollar e implementar computacionalmente un algoritmo que procese la información entregada por un radar, localizador GPS y sensor de azimuth digital, para estimar la localización geográfica absoluta de las orillas detectadas.
- Aplicar técnicas de procesamiento de señales para reducir la incertidumbre asociada a las coordenadas WGS84 obtenidas.
- Evaluar el desempeño de la solución implementada en función de la calidad de sus estimaciones frente a las técnicas de procesamiento de señales utilizadas en aplicaciones similares.

1.3 APORTES ESPERADOS

Al lograr esta tesis se conseguiría la presentación de una nueva técnica de georreferenciación remota de orillas fluviales, con aportes a nivel científico, tecnológico, en competitividad, productividad y medio ambiente de la sociedad dadas sus aplicaciones en navegación fluvial e hidrografía, como se detallará a continuación.

1.3.1 A nivel Científico - Tecnológico

Desde un punto de vista tecnológico, se contribuirá en el desarrollo y aplicación de un método novedoso para la obtención de información geográfica a distancia a partir del procesamiento digital de la información entregada por diversos equipos electrónicos. Así, se contribuirá a las alternativas existentes para realizar el levantamiento cartográfico de las orillas e islas de canales navegables, presentándose una transferencia de tecnología desde el área electrónica de procesamiento al campo de la hidrografía.

De acuerdo con esto, se esperan los siguientes beneficios:

- Generación de nuevo conocimiento al requerirse la aplicación de técnicas de procesamiento y estimación de señales, uno de los campos de trabajo del Grupo de Telecomunicaciones y Señales de la Universidad del Norte.
- Disponibilidad de resultados tangibles para la presentación de artículos científicos que permitan la apropiación social del conocimiento generado.
- Posibilidad de productos patentables en cuanto al método y sistema de georreferenciación a desarrollar en esta tesis.

1.3.2 A nivel de Productividad y Competitividad

El presente trabajo contribuirá de manera importante en la viabilidad de una propuesta de georreferenciación cuyos costos son reducidos en comparación a las metodologías existentes. La reducción de costos permitiría mejores rendimientos de las instituciones que implementen estos procesos, además se contribuye al mantenimiento y ampliación de los sistemas de navegación electrónica para vías fluviales.

Dicho pronóstico se realiza considerando la demanda proyectada para el sistema como las necesidades de georreferenciación de tramos fluviales que surjan en el país medida en kilómetros. Los niveles de ingreso dependerán de las tasas internas de retornos que cada entidad que lo implemente decida. A nivel de región Caribe, se espera que la técnica desarrollada desplace a las tecnologías que se vienen utilizando para georreferenciación de orillas para cartas náuticas fluviales. Particularmente, se proyecta una constante utilización en el mantenimiento del Sistema de Navegación Satelital, lo que garantiza los 1.000 Km. entre Puerto Salgar y Barranquilla.

Adicionalmente, la aparición de una técnica de georreferenciación a costos reducidos le permitirá al estado Colombiano aumentar la frecuencia de actualización del “SNS” a las adecuadas de acuerdo a los requerimientos de sectores trenzados y meándricos¹ en el sector, según se presenta en la Tabla 1.1. Siguiendo esta directiva, la demanda de georreferenciación por parte del SNS será de 3.959 Km.

Tabla 1.1 Frecuencias de Actualización recomendada para sectores entre Barranquilla (K0)-Puerto Salgar (K921)

Sector	Característica	Frecuencia Recomendada	Km/año
Barranquilla (K0)-Capulco (K503)	Meándrico	Anual	503
Capulco (503)-Berrío(K765)	Trenzado	Mensual	3144
Berrío(K765)-Puerto Salgar(K921)	Mixto	Semestral	312

Fuente: (Alvarado, 2009)

De manera similar, a nivel nacional, Colombia cuenta con varios ríos navegables los cuales presentan longitudes relacionadas en la Tabla 1.2, los cuales ampliarían el mercado de georreferenciación siempre que el Estado Colombiano le apueste a la señalización y recuperación de sus vías fluviales.

Tabla 1.2. Potencial requerimiento de georreferenciación en ríos de Colombia.

Vía Fluvial	Longitud aprox.
Río Atrato	1.000 Km.
Río Meta	700 Km.
Río Guaviare	1.400 Km.
Río Putumayo	1.500 Km.

Fuente: Ministerio de Transporte (Colombia).
Manual de Ríos Navegables.

1.3.3 En el medio Ambiente y la Sociedad

Al contribuir en un proceso de georreferenciación de orillas e islas de ríos, se posibilita una mayor disponibilidad de datos geográficos útiles en estudios hidrográficos y de morfología fluvial, control de inundaciones y planeación de obras civiles en busca de la mejora en la estabilidad de los canales navegables fluviales. El potenciamiento de las mencionadas actividades, sin duda que influirá en un mayor y mejor aprovechamiento de la riqueza hídrica del país.

¹ Tipo de morfología fluvial que consiste en un solo canal a través de sectores curvos (meandros) y rectos.

2 MARCO REFERENCIAL DE LA INVESTIGACIÓN

2.1 MARCO TEÓRICO

El área de estudio del presente proyecto de investigación se enmarca en el campo de la ingeniería electrónica en cuanto el problema científico, la propuesta de solución esbozada, el proceso de recolección de información de campo, así como el tratamiento de las variables de entrada se realiza mediante implementaciones de hardware y técnicas computacionales de esta disciplina. Es decir, que el nivel de desarrollo requerido en ésta tesis para dicha área, la transforma desde su categoría de medio a objeto de estudio teniendo en cuenta que en un contexto interdisciplinar como la georreferenciación, generalmente la electrónica representa un recurso o medio para obtener, procesar y almacenar la información.

2.1.1 Acerca de la Georreferenciación

Es común en la bibliografía, que se defina la georreferenciación como un proceso aplicado a imágenes, dada la gran cantidad de aplicaciones que parten del procesamiento de estas para obtener información geográfica (Reuter, 2005). Sin embargo, en este trabajo se utilizará un criterio más amplio definiéndolo como un proceso en el cual a un objeto físico, se le asigna una característica unívoca que es su posición geográfica con respecto a un sistema de referencia dado. Luego, si el proceso se aplica a conjuntos de objetos clasificados en determinadas categorías, la organización de todos estos elementos puede conformar un Sistema de Información Geográfico (SIG).

Por ejemplo, si se realiza la georreferenciación de los árboles de la ciudad, además de su clasificación, es posible almacenar esta información en una base de datos, que permita mostrar la distribución geográfica de estos por especies, tamaños, edad, estado de salud, etc. El mismo trabajo se podrá realizar con objetos como edificaciones, construcciones, carreteras, lagos, ríos, profundidades del río, personas o vehículos, etc, siendo que para estos dos últimos, dado su carácter móvil, se debe proveer un medio de comunicación para transmitir sus coordenadas al centro de datos.

El propósito de los SIG es brindar un ambiente apropiado que permita el análisis correcto y sencillo de los datos geográficos, proporcionando así una herramienta óptima capaz de ayudar en el proceso de toma de decisiones en aquellos casos que la componente geográfica sea de gran interés (ESRI, 2000). En la actualidad existen diversas formas de modelar dichas relaciones. Sin embargo, muchas de estas son complejas; los SIG facilitan este proceso al ser implementados mediante archivos raster y vectoriales.

2.1.2 Acerca del Sensado Remoto

El sensado remoto comprende un conjunto de técnicas que permiten la obtención de información de un objetivo por parte de un dispositivo separado en alguna distancia. Si bien bajo esta definición se abarca a la gran cantidad de aplicaciones civiles y militares que han hecho uso de la fotografía, ondas de radio y ultrasonidos, es necesario precisar que el término empezó a ser utilizado desde los años 60 para referirse a la utilización de sensores en todo el espectro electromagnético para medir variables de interés desde satélites y aviones en aplicaciones como las listadas en la Tabla 2.1. La utilización de embarcaciones también está contenida en el sensado remoto con aplicaciones en hidrografía.

Tabla 2.1. Algunas áreas de aplicación del sensado remoto

Área	Aplicación de Sensado Remoto
Cartografía	Reconocimiento y levantamiento de la geografía
Geología	Levantamientos, recursos minerales
Uso de la tierra	Urbanización, Agricultura, levantamiento de suelos, salud de cultivos, humedad y evatranspiración de suelos, seguimiento a lluvias acidas, deforestación, distribución de fauna y flora
Ingeniería Civil	Estudio de sitios, Recursos hídricos, transporte
Recursos Hídricos	Superficies de agua, fuentes, polución bajo agua, reconocimiento y levantamiento de glaciales y témpanos
Estudio de Costas	Erosión, batimetrías, monitoreo de emisiones térmicas y polución química
Oceanografía	Temperatura de superficie, estudio del geoide, topografía del fondo, corrientes, oleaje y mareas, circulación, monitoreo de combustibles
Meteorología	Seguimiento al estado del tiempo, pronóstico, flujo de calor, balance de energía, modelos de circulación, monitoreo de precipitaciones, procesos de desertificación,
Desastres Naturales	Inundaciones, terremotos, erupciones volcánicas, incendios forestales, incendios subterráneos de carbón, deslizamientos, tsunamis
Astronomía	Todas las anteriores aplicadas al exterior del planeta

Fuente: Cracknell y Hayes, 2007.

Algunos sistemas de sensado remoto se listan a continuación:

- CASI (Compact Airborne Spectral Imager)
- LANDSAT
- LIDAR (Light Detection and Ranging)
- SPOT
- RADARSAT (Sistema Satelital para Tomas de Radar)
- IRS (Indian Remote Sensing Satellite)
- IKONOS
- QUICKBIRD
- ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer)

- MOMS (Modular Optoelectronic Multispectral Stereoscopic-Scanner)
- CBERS (WFI, HRCCD, IR-MSS)
- MODIS (Moderate-resolution Imaging Spectroradiometer)
- MOS (Modular Optoelectronic Scanner).

La selección entre un sensor remoto aéreo y uno satelital depende, entre otros específicos de la aplicación, de los siguientes factores:

- Extensión del área cubierta
- Velocidad de desarrollo del fenómeno observado
- Nivel de detalle requerido y disponible por la tecnología en cada caso.
- Disponibilidad y costo de los datos.

El último factor, que tiende a ser minimizado en la discusión científica, termina siendo determinante en la selección de la tecnología de sensado remoto. Para el caso de los satélites, la información detallada y específica proporcionada por el dispositivo, estuvo durante mucho tiempo solo al alcance de entidades gubernamentales de los países propietarios y sus aliados. Recientemente, debido a cambios en las políticas de acceso, existen compañías que comercializan la información satelital, como los servicios del IKONOS y QUICKBIRD que ofrece fotografías con resoluciones menores a un metro. Los costos y tiempos de acceso determinan que para algunos países la utilización sea prácticamente inviable.

2.2 MARCO CONCEPTUAL

2.2.1 Representación de Sistemas de Información Geográfica (SIG)

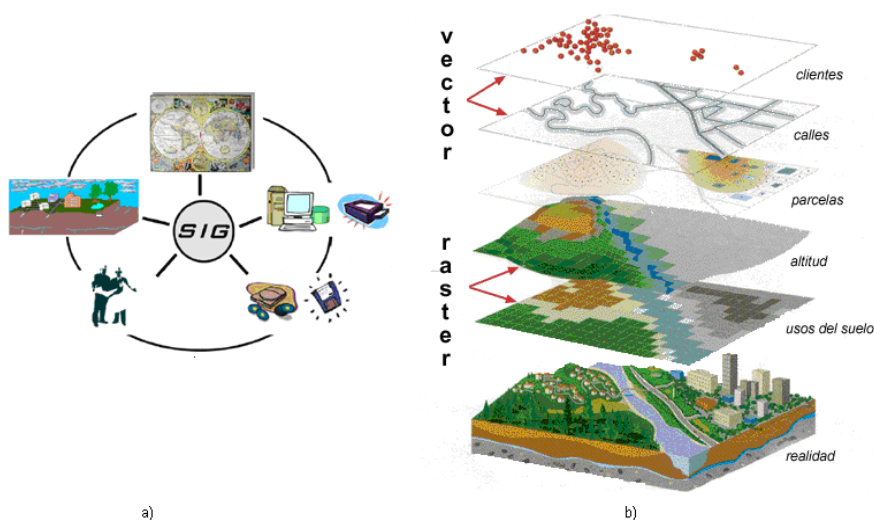
Los SIG simplifican la presentación de la información temática y geográfica mediante la utilización combinada de dos tipos de archivos: raster y vectorial.

2.2.1.2 Archivo raster: Los sistemas de información raster basan su funcionalidad en el principio de las relaciones de vecindad entre los objetos geográficos. Este método relaciona la matriz de píxeles asociada a una imagen y atribuye un valor numérico a cada celda como representación de su valor temático. Dado que los tamaños de los píxeles son constantes, y que además, se conoce la posición en coordenadas cartesianas tridimensionales del centro de una de las celdas, se puede decir que todos los píxeles se encuentran georreferenciados (ver Figura 2.1a y 2.1b).

2.2.1.1 Archivo vectorial: Son aquellos que para la descripción de los objetos geográficos utilizan vectores definidos por pares de coordenadas relativas a algún sistema de referencia. Con un par de coordenadas y su altitud crean un punto, con dos puntos generan una línea, y con un conjunto de líneas forman polígonos. (Ver figura 2.1b). La información geográfica así almacenada puede tener un nivel de detalle elevado manteniendo tamaños de almacenamiento reducidos. Algunos ejemplos de este tipo de

archivos son los .kml utilizados en el software Google Earth para almacenar puntos de interés, rutas y marcaciones por parte del usuario, los .shp estandarizados por ARGIS para la conformación de Sistemas de Información Geográfica.

Figura 2.1. a) Principales componentes de un sistema de información geográfica.
b) Visualización de los SIG Vectoriales y Raster.



Tomado de (ESRI, 2000).

La información a generar a través de la propuesta de esta investigación consistirá en la secuencia de puntos de las orillas fluviales, lo que constituye la base para que mediante algún software editor de información geográfica (ARGIS, AUTOCAD MAP, etc) se obtenga bajo la presentación y/o estándar requerido.

2.2.2 Sistemas de Referencia Geográfica

Un sistema de referencia es el conjunto de convenciones y conceptos teóricos que definen en cualquier momento la orientación, ubicación de los objetos en el espacio con respecto a tres ejes coordenados (X , Y , Z) (IGAC, 2004). Si un sistema de referencia tiene su centro en coincidencia con el centro terrestre, se denominará geocéntrico. La adopción de un sistema de referencia geocéntrico está íntimamente ligada a la escogencia de un elipsoide de referencia, el cual no es más que un modelo idealizado del geoide terrestre. La dificultad para determinar la ubicación exacta del centro de la tierra y la necesidad de cada país en contar con un modelo elipsoidal que se ajustara mejor a su topografía, generó las apariciones de datums o sistemas de referencia relacionadas con un elipsoide.

Actualmente, el datum de referencia más aceptado es el WGS-84 que modela a la tierra de acuerdo con el elipsoide cuyos parámetros se presentan en la Tabla 2.2.

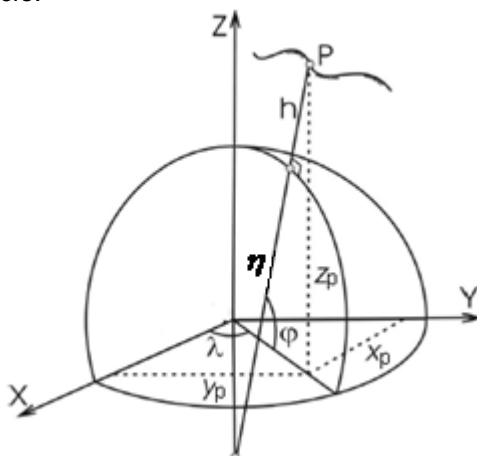
Tabla 2.2. Constantes del elipsoide WGS-84

Constante	Relación Matemática	Valor
Semieje mayor(a)	Constante definida	$a=6.378.137 \text{ m}$
Constante gravitacional geocéntrica (GM)	Constante definida	$GM=3.896.005 \times 10^8 \text{ m}^3 \text{ s}^{-2}$
Factor de aplanamiento dinámico (J_2)	Constante definida	$J_2=108.263 \times 10^{-8}$
Velocidad angular de rotación (ω)	Constante definida	$\omega = 7.292.115 \times 10^{11} \text{ rad.s}^{-1}$
Primera excentricidad (e^2)	Cálculo iterativo a partir de GM, J_2 y ω ,	$e^2=6,69438002290 \times 10^{-3}$
Segunda excentricidad(e'^2)	$e'^2 = \frac{e^2}{1 - e^2}$	$e'^2=6,73949677548 \times 10^{-3}$
Semieje menor(b)	$b = a\sqrt{1 - e^2}$	$b=6.356.752,31414 \text{ m}$
Aplanamiento recíproco($1/f$)	$f = \frac{a - b}{a}$	$1/f=298,257222101$

Fuente: IGAC, 2004.

La información geográfica en este sistema puede ser representada mediante coordenadas cartesianas tridimensionales o mediante coordenadas elipsoidales. Las coordenadas cartesianas tridimensionales expresan la ubicación geográfica de un punto estableciendo las distancias lineales entre la posición marcada en el sistema de referencia y cada uno de los planos del sistema coordinado. Las coordenadas son dadas en tríadas (x_p, y_p, z_p) y calculadas mediante el trazado de recta paralelas a los ejes desde el punto de interés hasta los ejes de referencia, siendo el tamaño en metros de cada una de estas rectas, equivalente a la coordenada cartesiana correspondiente (ver Figura 2.2).

Figura 2.2. Coordenadas cartesianas tridimensionales (x_p, y_p, z_p) y elipsoidales (λ, ϕ, h) de un punto en el espacio.



Tomado de (IGAC, 2004).

Las coordenadas elipsoidales son representadas mediante las cantidades angulares longitud (λ), latitud (φ) y altura elipsoidal (h), donde las dos primeras son ángulos en el sistema sexagesimal; mientras que la altura elipsoidal es expresada en metros y equivale a la distancia a lo largo de la gran normal (η) que hay entre la superficie del elipsoide de referencia y el punto a calcular. La latitud se define como el ángulo formado por el plano ecuatorial y la gran normal al elipsoide en el punto de interés, siendo positiva al norte o por encima del ecuador y negativa al sur, su rango de variación es de $-90 \leq \varphi \leq 90$; La longitud representa el ángulo sobre el plano ecuatorial entre el meridiano de origen o de Greenwich y el meridiano del punto a referenciar; variando desde $180 \leq \lambda \leq 180$ de oeste a este del Greenwich (Ver Figura 2.2).

Las coordenadas cartesianas y elipsoidales, pueden transformarse entre sí mediante deducción trigonométrica y la aplicación de los parámetros del elipsoide correspondiente (por ejemplo los de la Tabla 2.2), generándose las siguientes ecuaciones (IGAC, 2004):

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} (\eta + h) \cos \varphi \cos \lambda \\ (\eta + h) \cos \varphi \sin \lambda \\ ((1 - e^2) \eta + h) \sin \varphi \end{bmatrix} \quad (2.2.1)$$

$$\begin{bmatrix} \varphi \\ \lambda \\ h \end{bmatrix} = \begin{bmatrix} \arctan \left(\frac{z + e^2 b \sin^3 \vartheta}{\sqrt{x^2 + y^2} - e^2 a \cos^3 \vartheta} \right) \\ \arctan(y/x) \\ \frac{\sqrt{x^2 + y^2}}{\cos \varphi} - \eta \end{bmatrix} \quad (2.2.2)$$

Donde η y ϑ que se utilizan como variables auxiliares para efectos de simplificación de (2.2.1) y (2.2.2) se calcularán como:

$$\eta = \frac{a}{\sqrt{1 - e^2 \sin^2 \varphi}} \quad (2.2.3)$$

$$\vartheta = \arctan \left(\frac{za}{\sqrt{x^2 + y^2} b} \right) \quad (2.2.4)$$

2.2.3 Estimación Bayesiana de Parámetros

El proceso de estimación permite obtener el valor de un parámetro relacionado con las características estadísticas de un número finito de muestras. El parámetro a estimar puede ser aleatorio o no aleatorio. La estimación de parámetros aleatorios es conocida como Estimación de Bayes, mientras que la estimación de parámetros no aleatorios se

denomina Estimación de Máxima Verosimilitud MLE² (Mourant, 2005). Sin embargo, según la perspectiva presentada en (Vasegui, 2000) la MLE está contenida como variante dentro de la Estimación de Bayes.

Sean y_1, y_2, \dots, y_K las K independientes e idénticamente distribuidas muestras de una variable aleatoria Y , con cierta función de densidad de probabilidad que depende de un parámetro desconocido Θ . Se puede definir $g(y_1, y_2, \dots, y_K)$ como una función o estadístico para estimar el parámetro Θ , de tal forma que el estimador esté dado por:

$$\hat{\Theta} = g(y_1, y_2, \dots, y_K) \quad (2.2.1)$$

Siendo $g()$ calculada en procura de minimizar la función de riesgo dada por:

$$\mathfrak{R}(\hat{\Theta} | y) = \int_{\Theta} C(\hat{\Theta}, \Theta) \cdot f_{\Theta|Y}(\Theta | y) \cdot d\Theta \quad (2.2.2)$$

Donde $C(\hat{\Theta}, \Theta)$ es una función de costo para la solución obtenida y $f_{\Theta|Y}(\Theta | y)$ expresa la dependencia de la calidad de la estimación con el conocimiento estadístico a priori del fenómeno y las muestras de información que se disponen efectivamente. La aplicación del concepto de estimación y regla de Bayes sobre (2.2.2) conlleva a la expresión:

$$\hat{\Theta}_{Bayesian} = \arg_{\hat{\Theta}} \min \left[\int_{\Theta} C(\hat{\Theta}, \Theta) \cdot f_{Y|\Theta}(y | \Theta) \cdot f_{\Theta}(\Theta) \cdot d\Theta \right] \quad (2.2.3)$$

Dependiendo del criterio utilizado para implementar (2.2.3), se puede hablar de los siguientes tipos de estimación.

2.2.3.1 Estimación de Máximo a Posteriori (MAP):

Se busca maximizar la función de distribución de probabilidad posterior $f_{\Theta|Y}(\Theta | y)$, mediante la definición de la función de costo:

$$C_{MAP}(\hat{\Theta}, \Theta) = 1 - \delta(\hat{\Theta}, \Theta) \quad (2.2.4)$$

Donde $\delta(\hat{\Theta}, \Theta)$ es la función delta Kronecker. Al sustituir (2.2.4) en (2.2.2) y proceder a la disminución de la función de riesgo, el estimador bayesiano bajo el criterio de máxima probabilidad posterior se obtendrá como:

$$\hat{\Theta}_{MAP} = \arg_{\hat{\Theta}} \max [f_{Y|\Theta}(y | \Theta) \cdot f_{\Theta}(\Theta)] \quad (2.2.5)$$

² **MLE:** *Maximun Likelihood Estimation.*

2.2.3.2 Estimación de Máxima Verosimilitud (ML):

Se procura la obtención de un parámetro estimado que maximice la función de verosimilitud $f_{Y|\Theta}(y|\Theta)$ para lo cual se plantea la función de riesgo:

$$\begin{aligned}\mathfrak{R}_{ML}(\hat{\Theta} | y) &= \int_{\Theta} [1 - \delta(\hat{\Theta}, \Theta)] \cdot f_{Y|\Theta}(y | \Theta) \cdot f_{\Theta}(\Theta) \cdot d\Theta \\ &= \text{const} \cdot [1 - f_{Y|\Theta}(y | \Theta)]\end{aligned}\quad (2.2.6)$$

Donde se ha tomado la función de distribución a priori del parámetro desconocido Θ como una constante. Luego, la minimización de la función de riesgo se expresará como la maximización de la función de verosimilitud, por tanto:

$$\hat{\Theta}_{ML} = \arg_{\hat{\Theta}} \max [f_{Y|\Theta}(y | \Theta)] \quad (2.2.7)$$

La ecuación (2.2.7) puede ser resuelta mediante técnicas clásicas de cálculo, por lo que es válida la utilización del logaritmo natural para facilitar la solución, transformando la ecuación de verosimilitud en:

$$\hat{\Theta}_{ML} = \arg_{\hat{\Theta}} \text{zero} \left[\frac{\partial}{\partial \hat{\Theta}} \ln [f_{Y|\Theta}(y | \Theta)] \right] \quad (2.2.8)$$

Si $\hat{\Theta}$ cumple con (2.2.8) es un estimador de máxima verosimilitud para Θ . Por ejemplo, si Y tiene una función de probabilidad normal caracterizada por los parámetros media y varianza (μ, σ^2) , las estimaciones de máxima verosimilitud de estos estarán dadas por la media y varianza muestrales (Mourant, 2005):

$$\hat{\mu}_{ML} = \frac{1}{K} \sum_{k=1}^K y_k \quad (2.2.9)$$

$$\hat{\sigma}_{ML}^2 = \frac{1}{K} \sum_{k=1}^K (y_k - \mu)^2 \quad (2.2.10)$$

2.2.3.3 Estimación de Mínimo Error Cuadrático Medio (MMSE):

Se obtiene con el parámetro que minimice el error cuadrático medio del error, de tal forma que la función de riesgo esté dada por:

$$\mathfrak{R}_{MMSE}(\hat{\Theta} | y) = \int_{\Theta} (\hat{\Theta} - \Theta)^2 \cdot f_{\Theta|Y}(\Theta | y) \cdot d\Theta \quad (2.1.11)$$

Luego la minimización de (2.1.11) genera como solución:

$$\hat{\Theta}_{MMSE} = \int_{\Theta} f_{\Theta|Y}(\Theta | y) d\Theta \quad (2.2.12)$$

Para situaciones en que no existe conocimiento a priori de la estadística de Θ , el criterio se transforma al menor error cuadrático (e), quedando el estimador como:

$$\hat{\Theta}_{LSE} = \arg_{\hat{\Theta}} \min [E[e^2(\Theta | y)]] \quad (2.2.13)$$

Siendo el operador $E[\cdot]$ la esperanza matemática. La versión de estimación por reducción del error cuadrado dado en (2.2.13) permite el cálculo sin conocimiento a priori de las distribuciones estadísticas del parámetro o las muestras del proceso.

2.2.3.4 Estimación de Mínimo Error Absoluto Medio (MAVE):

Procura la reducción del error absoluto medio definiendo la función de riesgo como:

$$\mathfrak{R}_{MAVE}(\hat{\Theta} | y) = \int_{\Theta} |\hat{\Theta} - \Theta| \cdot f_{\Theta|Y}(\Theta | y) \cdot d\Theta \quad (2.2.14)$$

Al aplicar la definición de valor absoluto y minimización de la función de riesgo:

$$\hat{\Theta}_{MAVE} = \arg_{\hat{\Theta}} \text{zero} \left[\frac{\partial}{\partial \hat{\Theta}} \mathfrak{R}_{MAVE}(\hat{\Theta} | y) \right] \quad (2.2.15)$$

$$\hat{\Theta}_{MAVE} = \arg_{\hat{\Theta}} \text{zero} \left[\int_{-\infty}^{\hat{\Theta}} f_{\Theta|Y}(\Theta | y) + \int_{\hat{\Theta}}^{\infty} f_{\Theta|Y}(\Theta | y) \right] \quad (2.2.16)$$

De la ecuación (2.2.16) se debe notar que $\hat{\Theta}_{MAVE}$ se corresponde con la mediana de la función de probabilidad posterior de las muestras (Vasegui, 2000).

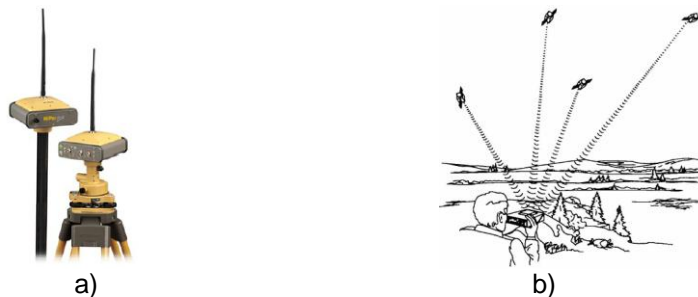
2.2.4 GPS: Nociones Básicas de Operación

El sistema GPS (*Global Positioning System*, por sus siglas en inglés) permite determinar las coordenadas de posición en cualquier punto de la superficie terrestre, a partir del procesamiento de señales emitidas por satélites orbitando en el espacio exterior.

Su funcionamiento se basa en una constelación de veintinueve satélites que giran en torno a la tierra describiendo órbitas de aproximadamente 20.000km (no geostacionarias). Cada uno de estos satélites se encarga de enviar constantemente ondas de radio que contienen información relacionada con su posición actual mediante la

incorporación de tiempos de envío en estas señales. El receptor ubicado en la superficie de la tierra, debe estar en la posibilidad de captar señales de por lo menos cuatro de los satélites disponibles para triangular su posición (Véase Figura 2.3b). Una vez se cumple esta condición, el receptor intenta estimar la distancia de separación entre este y los satélites mediante el uso de la información horaria incluida en las señales de radio capturadas. Resulta de vital importancia la sincronización, precisión y exactitud para estos cálculos de tiempo; en este sentido todos los satélites utilizados en esta aplicación cuentan con relojes atómicos de alta precisión y los receptores terrestres utilizan relojes avanzados que les permiten realizar cuentas de tiempo en el orden de los nanosegundos.

Figura 2.3. a) GPS TopCon Híper Pro. b) Localización de un punto de la tierra mediante el sistema GPS



Tomado de (Pozo, 2005) y (TOPCON, 2010).

El cálculo de la posición es muy vulnerable a errores principalmente por la dependencia que posee el sistema al tiempo. Una de las variables más utilizadas para describir y evaluar errores en los cálculos llevados a cabo por el receptor es la dilución de precisión en posición (PDOP, por sus siglas en inglés), entre más pequeña sea esta figura de mérito más alineados estarán los satélites al receptor de manera perpendicular dificultando el cálculo de la posición. Además de esta fuente de error geométrico, las señales de radiofrecuencia enviadas por los satélites son vulnerables a cambios en su velocidad de propagación debido a partículas cargadas ubicadas en la ionosfera y al vapor de agua de la troposfera. Factores como el multi-trayecto de las señales, la calidad del equipo y la disponibilidad selectiva (restricción de origen militar) juegan también un papel importante a la hora de discriminar y minimizar errores.

Por los anteriores problemas, las mediciones de posición realizadas traen asociado un cierto grado de incertidumbre, entregando una estimación o valor medio con una determinada dispersión. El error cometido es generalmente caracterizado mediante el error cuadrático medio (ECM) definido como la raíz cuadrada de la media de los errores al cuadrado, pudiendo referirse a una, dos o tres dimensiones. En los receptores GPS, los errores siguen una distribución de probabilidad normal en cada eje, permitiendo deducir las probabilidades asociadas a los mismos de una manera sencilla. Los fabricantes generalmente definen la precisión de las medidas con el Error Circular Probable (ECP), que equivale al radio del círculo en donde se encuentra la estimación más probable con un 95% de confiabilidad.

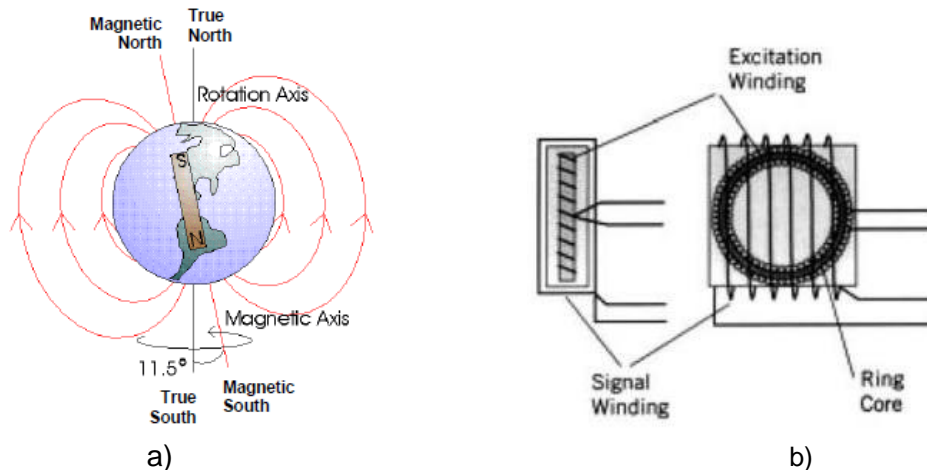
La información de la posición entregada por un dispositivo GPS incluye la latitud, longitud y altura elipsoidal, en el formato sexagesimal, es decir, en grados, minutos y segundos referidos al elipsoide WGS84, además de la fecha y hora exacta en que la medición fue realizada. Generalmente los datos son proporcionados en protocolo RS-232 bajo el estándar NMEA 0183.

2.2.5 Sensor de Azimut Electrónico: Nociones Básicas de Operación

La naturaleza de operación de estos sensores varía en función del instrumento de medición interno. De manera específica, en la implementación de esta tesis se utilizó un sensor de azimut electrónico Furuno C500 basada en un magnetómetro *fluxgate*. En la Figura 2.4a se muestra la forma del campo magnético generado por la tierra. El norte marcado por la forma y dirección del campo no coincide con el norte geográfico determinado por el eje de rotación de la tierra, desviándose en función de la geografía terrestre. La medida arrojada por un sensor de Azimut se basa en su capacidad de medir la desviación angular con respecto al norte magnético mediante la variación del campo magnético terrestre medido. El magnetómetro (Véase Figura 2.4b) transforma estas variaciones magnéticas en señales eléctricas donde un sistema electrónico digital la procesa para transformarla en datos de azimut en tiempo real, que serán transmitidos por cable mediante un estándar serial como el NMEA.

El rango de medición de estos dispositivos va desde 10^{-4} a 5 mT pudiendo alcanzar una resolución de 0.1nT, lo que lleva a resoluciones angulares como el caso del dispositivo utilizado de $\pm 0.1^\circ$.

Figura 2.4. a) Campo Magnético Terrestre. b) Estructura interna de un magnetómetro *fluxgate*



Tomado de (Phillips Semiconductor, 2000).

2.2.6 Radar: Nociones Básicas de Operación

El término radar proviene del acrónimo en inglés *Radio Detection And Ranging* (Detección y medición de rango por Radio), dicho sistema basa su funcionamiento en el fenómeno de reflexión de las ondas electromagnéticas en objetos sólidos y superficies. Estructuralmente consiste en una antena emisora/receptora de señales de radio que permite detectar obstáculos o volúmenes específicos en un espacio determinado mediante la medición de la intensidad de la onda reflejada en dicho obstáculo, conocida como eco, aprovechando que las ondas irradiadas presentan poca atenuación bajo diversas condiciones atmosféricas tales como la neblina o la lluvia. La ecuación matemática que resume su concepto de detección y rango es la siguiente:

$$R = c \cdot t_0 / 2 \quad (2.2.17)$$

Donde R es la distancia al obstáculo detectado, c es la velocidad de la luz en el medio, y t_0 es el retardo de tiempo que le toma al pulso enviado alcanzar el objetivo y regresar a la antena. Una vez el pulso modulado es irradiado por la antena al espacio debe transcurrir el tiempo suficiente que permita a las señales eco regresar a su fuente para ser detectadas, antes que sea enviado el siguiente pulso. Este tiempo es determinado en función de la potencia de radiación del equipo y del rango máximo de medición requerido.

Para el caso de radares marinos la antena emisora/receptora es mantenida sobre un eje giratorio que le permite alcanzar a cubrir completamente su alrededor. La velocidad de giro de la antena está estrechamente relacionada con factores mecánicos y factores ligados al procesamiento de los pulsos que son enviados desde esta y adicionalmente al ángulo de apertura del patrón de radiación de la onda electromagnética enviada en una dirección de giro fija. El nivel de apertura de este ángulo es el ancho del patrón de radiación de la antena, que definirá que tanta área del objetivo es observada por el radar, y por ende los valores de rango y resolución para cada pulso enviado por el radar analizado en un ámbito espacial. El ancho de haz B para un alcance R estará dado por [33]:

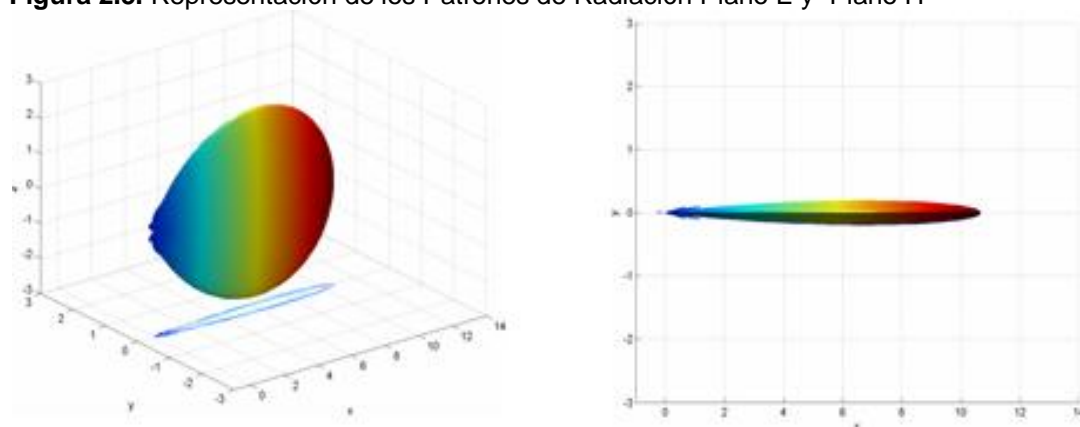
$$B = 2 \cdot R \cdot \tan(\phi/2) \quad (2.2.18)$$

Donde ϕ es el ancho horizontal del patrón de radiación, el cual generalmente tiene la forma mostrada en la Figura 2.5.

Una vez el pulso de onda es transmitido por la antena del radar este es susceptible a golpear cualquier objeto eléctricamente visible (con proporciones físicas mayores a la longitud de onda de la señal irradiada), generando un grupo de ondas dispersas con dirección variable. Las direcciones, potencia y los ángulos con los que se dispersa la radiación incidente, y aún más importante, el porcentaje la potencia de radiación que retorna en la misma dirección de la onda incidente dependen tanto de las propiedades físicas del objeto (forma, material, velocidad, etc.) como de las propiedades de la onda

irradiada y de su fuente, existiendo para cada forma geométrica un patrón de retro dispersión asociado. Dicho porcentaje correspondiente a la porción de potencia que regresa en la misma dirección y con la misma polarización que la onda transmitida permite definir la Sección radar equivalente (RCS en inglés, Radar Cross Section) del objetivo. La RCS es una propiedad del objetivo definida como el área proyectada de una esfera de metal que retornaría la misma señal eco si esta fuese intercambiada con el objetivo. Matemáticamente se puede definir la RCS mediante la relación entre la densidad de potencia transmitida y recibida de un objetivo.

Figura 2.5. Representación de los Patrones de Radiación Plano E y Plano H



Tomado de (Ruiz y Villar, 2007).

Sin embargo, la radiación que retorna a la antena del radar no corresponde sólo a la RCS del objetivo de interés (en caso que el radar sea usado bajo la perspectiva de detección y rastreo de objetos), también contiene información relacionada con el entorno físico que rodea a este (suelo, árboles, mar, etc.). A aquellas señales detectadas por el radar que corresponden a la radiación retorno de volúmenes y superficies que no corresponden al objeto de interés se le conoce como *clutter* y en el modo de operación tradicional de un radar puede interferir con la medición y capacidad de detección de este. De manera general existen dos tipos de *clutter*: superficial (suelo, mar, vegetación, formaciones rocosas, etc.) y volumétrico (objetos aéreos principalmente: aves, lluvia, insectos, etc.).

Si el clutter es intenso y no filtrable, termina siendo sintetizado como un objetivo por el radar, conociéndose como blanco disperso. La superficie de orillas e islas corresponden con este modelo de objetivo, de ahí que su localización presente una incertidumbre mayor a la intrínseca del radar.

2.2.7 Detección de Líneas y Bordes en Imágenes

En el capítulo 3 se mostrará que el método de adquisición optado para la información de radar proporciona imágenes de la pantalla del dispositivo, se requiere la utilización de técnicas de detección de bordes en imágenes, lo que permitirá obtener la ubicación de las orillas trazadas por el radar.

Un borde es la frontera entre dos regiones con propiedades de nivel de gris distintas. Para la detección de bordes, una técnica de fácil implementación es el cálculo del gradiente del color (tasa de cambio) mediante el producto de sectores de pixeles alrededor de cada punto con máscaras o pequeñas cuadrículas de pixeles como la definida en la Tabla 2.3., donde las componentes $z_1 \dots z_9$ definen posiciones en una máscara de 3x3.

Tabla 2.3. Asignación de Pixeles para Máscara de Exploración de Imágenes de Radar

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

Los bordes se pueden obtener aplicando un operador de gradiente de Sobel, definidos en (González y Woods, 1996):

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3) \quad (2.2.19)$$

$$G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7) \quad (2.2.20)$$

Luego, debido a que las imágenes a procesar en este trabajo son prácticamente binarias, la cantidad $G_y + G_x$ se revisará con respecto a un umbral para tomar la decisión de si un determinado pixel se encuentra sobre un borde.

2.3 ESTADO DE DESARROLLO DEL TEMA

2.3.1 Trabajos y Técnicas Actuales en Georreferenciación y Sensado Remoto Aplicado

En una carta náutica, el trabajo consiste en identificar mediante coordenadas geográficas la línea de orilla del canal principal y las islas, a la que se añade la línea de navegación que se actualiza constantemente mediante trabajos de batimetría. La referencia geográfica del contorno de orillas e islas se toma durante el estiaje, niveles mínimos del río cuando es posible definir totalmente las islas, playones y otros accidentes que no serían visibles durante niveles medios y altos. En estas condiciones, la obtención de coordenadas geográficas para generar una carta digital de un tramo en estudio es posible mediante técnicas utilizadas en la conformación Sistemas de Información Geográficos, tales como (Cracknell y Hayes, 2009):

- Fotografía aérea digital Georreferenciada.
- Fotografía satelital Georreferenciada.
- RADARSAT

- LIDAR.
- Levantamiento topográfico directo con GPS de alta precisión.

Una aplicación del sensado remoto satelital que debe destacarse es el sistema Google Earth que integra información de tomas satelitales y otras fuentes para proveer a sus usuarios de una vista georreferenciada del mundo, con muy alto nivel de detalle. El costo del proyecto es subsanado por las ventas de publicidad y posicionamiento comercial de GOOGLE. La vigencia de la información es aleatoria, con buena actualización de los principales entornos urbanos y alta caducidad en zonas apartadas y rurales, como la mayoría de las riberas de los ríos colombianos.

En el caso del sensado remoto aéreo, las dificultades de acceso vienen dadas por el alto costo de la tecnología de sensado y de los vuelos particulares requeridos para la recolección de información, además del alto nivel de “know how” de cada técnica en particular. El acceso a la información ya procesada es provisto mediante contratación a compañías especializadas proveedoras de estos servicios.

De estas tecnologías, generalmente los usuarios escogen como alternativas aquellas que cumplan los requerimientos de resolución, exactitud planimétrica y suministre la información requerida en la aplicación dada. Así se encontraron dentro del campo de aplicación los siguientes trabajos relacionados:

- En el trabajo “Sistema de Asistencia Satelital a la Navegación en el Río Magdalena entre Puerto Salgar (K925) y Barranquilla (SNS)” (Alvarado, 2009) para los requerimientos de señalización satelital en el SNS y dada las condiciones meteorológicas, se escogieron como alternativas las tecnologías de levantamientos topográficos directos de orillas e islas con GPS de alta precisión, en los puntos de control y segmentos reducidos, y restitución cartográfica mediante imágenes RADARSAT de alta resolución, en mayor proporción de terreno, requiriéndose posteriormente el proceso de ortorectificación y edición. En esta técnica se requiere trabajo de campo para la toma directa de coordenadas de ciertos puntos de control y gran cantidad de tiempo en actividad de campo debido a las grandes dimensiones del terreno y sus particularidades.
- En “Coastal dynamic and shoreline mapping: multi-sources spatial data analysis in Semarang Indonesia” (Marfai y otros, 2008) se reporta el seguimiento de la línea costera indonesia para los últimos 100 años mediante mapas topográficos antiguos (1908, 1937) y el análisis de imágenes satelitales LANDSAT e IKONOS, información que es georreferenciada para extraer conclusiones acerca de la movilidad del orilla.
- En “Assessment and integration of conventional, RTK-GPS and image-derived beach survey methods for daily to decadal coastal monitoring” (Harley y otros, 2010) se evalúa la validez de mediciones de la línea costera para las técnicas

RTK-GPS³, Video de alta resolución y levantamientos tradicionales realizados con postes de referencia para el seguimiento de la dinámica de estas orillas. Los autores logran validar los levantamientos con GPS como una alternativa de costo razonable al seguimiento con LIDAR y el video para seguimiento remoto con limitaciones a distancias medianas.

- El documento de patente “Method and Apparatus for Mapping and Measuring Land” (Ciampa, 1993), reporta la invención de un sistema de georreferenciación basado en sensor aéreo CCD (Cámara digital) con el que se miden y triangulan distancias desde los elementos de la geografía hasta el avión. Las tomas recopiladas y la información de distancias medidas posteriormente serían utilizadas para georreferenciación de los terrenos explorados.

La mayoría de trabajos reportados hacen especial énfasis en la exactitud de las técnicas, para aplicaciones generalmente distintas de navegación tales como geodesia, topografía, geología, con prevalencia del sensado remoto con tomas satelitales y aviones, dadas las dimensiones cubiertas. Ante los costos de obtención y utilización de la información de sensado remoto, se han gestado desarrollos que desvían el uso convencional de algunas tecnologías con la finalidad de darle uso en aplicaciones de interés científico, tal es el caso de los radares comerciales para navegación cuya señal contiene mayor información que la reportada por el dispositivo⁴. Se tienen trabajos como “Ocean Surface Determination from X-band radar-image sequences” (Dankert y Rosenthal, 2004) donde se presenta la adaptación de métodos de estimación de parámetros del mar por SAR y SLAR a mediciones de clutter marino obtenidas con radar común de navegación mediante tarjetas de adquisición acopladas a la señal bandabase de pulsos (video). En Colombia, se está buscando el desarrollo de esta tecnología, mediante el proyecto relatado en (Posada y Velez, 2010), el cual es financiado por la Universidad del Norte desde junio de 2010.

En el mismo sentido, de aprovechamiento de radares comerciales para mediciones, en el artículo “Bathymetry Determination via X-Band Radar Data: A New Strategy and Numerical Results” (Serafino y otros, 2010) se utilizan correlaciones entre el espectro de oleaje medido del mar y las profundidades de este en la correspondiente zona, planteándose batimetrías de poca profundidad. En “Retrieval of Surface-Current Fields and Bathymetries using Radar-Image Sequences” (Dankert, 2003) se proponen algoritmos para la determinación de batimetría y corrientes superficiales del mar, además de algún seguimiento de la dinámica de línea costera.

Teniendo en cuenta las líneas de investigación revisadas, no se encontraron trabajos que pretendan el cálculo o estimación de la posición geográfica de orillas mediante radar de navegación. En la presente investigación se desarrollan aspectos prácticos y teóricos necesarios para el levantamiento cartográfico de orillas fluviales mediante su sensado remoto con radar comercial.

³ Real Time Kinematics, sistema de monitoreo de movimiento tridimensional en tiempo real utilizando dos GPS de alta precisión (base y estación móvil).

⁴Distancias, azimut y velocidades de obstáculos alrededor de la embarcación.

2.3.2 Trabajos y Técnicas Actuales en Estimación de Parámetros y Variables

Es importante notar que en la sección 2.2.3 se ha presentado una reducción del concepto de estimación bayesiana a la búsqueda de parámetros, pero que en general Θ no es unidimensional, pudiendo representar en un momento dado el estado de un sistema o una señal. Por definición, la solución de (2.2.3) se resuelve como:

$$\hat{\Theta}_{Bayesian} = \arg_{\hat{\Theta}} \text{zero} \left[\frac{\partial}{\partial \hat{\Theta}} \int_{\Theta} C(\hat{\Theta}, \Theta) \cdot f_{Y|\Theta}(y | \Theta) \cdot f_{\Theta}(\Theta) \cdot d\Theta \right] \quad (2.3.1)$$

Las distintas variantes de estimación bayesiana expuestas en la sección 2.2.3 son utilizadas en prácticamente todos los campos de la experimentación científica, donde se requiera la determinación del modelo de un sistema con sus parámetros, de ahí que se mantenga vigente su utilización, como se presenta en el artículo “An introduction to Bayesian methods for analyzing chemistry data” (Armstrong y Hibbert, 2009), donde se retoman los fundamentos teóricos y el razonamiento de Bayes y se presenta paulatinamente su utilización como herramienta para el análisis de datos en química.

Sin embargo, los trabajos de investigación en el área actualmente están concentrados principalmente a la solución de dificultades con el tamaño de las muestras, la necesidad de censurar datos y la aplicación de Bayes a conjuntos de datos para los que no ha sido posible ajustar aceptablemente alguna función de distribución de probabilidad (modelos no paramétricos). En este sentido se pueden destacar los siguientes trabajos:

- En el artículo “Bayesian Robust Estimation of the Mean” (Wesley y Utts y Pearson, 1986) se propone un estimador para la media que reduce la influencia del sesgo ocasionado por una proporción de muestras contaminadas. La estimación aplica una corrección a la media muestral sumándole una cantidad calculada con una función de ponderación que depende del grado de dispersión de cada muestra en particular.
- En “Maximum likelihood estimation in semiparametric regression models with censored data” (Zeng y Lin, 2007) se concluye que la estimación por máxima verosimilitud es válida para modelos semi-paramétricos, desarrollando técnicas numéricas para el computo de las inferencias aún en sistemas de parámetros m-dimensionales.
- En el trabajo “Effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter of Poisson-gamma models for modeling motor vehicle crashes: A Bayesian perspective” (Dominique y Miranda, 2008) se estudia el efecto de la baja disponibilidad de muestras sobre la exactitud en el cálculo de la media en una distribución Poisson-Gamma utilizada en el modelado de fallas en motores vehiculares.

Además, las estrategias matemáticas y computacionales para solucionar (2.3.1) de manera eficiente y con datos disponibles recursivamente, derivó en algoritmos de

estimación (filtros) siendo los más conocidos Wiener, Kalman y EM, sobre los que se vienen haciendo importantes mejoras y críticas. A partir de esta área de conocimiento, se han consultado diversos trabajos relacionados con la estimación de señales y parámetros de éstas vía máxima verosimilitud y filtros. Las aplicaciones encontradas, de diversa índole, coinciden en la necesidad de estimar el estado de una variable mediante mediciones ruidosas buscando el óptimo mínimo error cuadrático. Las técnicas se basan en el modelo matemático de los sistemas y las covarianzas de ruido que los afectan. El rasgo característico de estas aplicaciones es la no ergodicidad del ruido filtrado, requiriéndose la adaptabilidad de los estimadores utilizados, funcionalidad que en algunos casos se apoya en técnicas de inteligencia artificial. Entre los trabajos revisados se pueden destacar:

- “Statistical recursive filtering for offset nonuniformity estimation in infrared focal plane array sensors” (San Martín, Torres y Pezoa, 2008) donde se propuso un filtro Kalman recursivo para determinar y compensar el factor de iluminación ambiental presente en cada píxel de imágenes tomadas con sensores infrarrojos.
- El trabajo “Adaptive tuning of Kalman filter via fuzzy logic for an intelligent AUV navigation system” (Loebis y otros, 2004) describe la implementación de un sistema de navegación inteligente basado en integración de INS (Inertial Navigation System) y GPS (Global Positioning System) mediante filtros de Kalman simple y filtro de Kalman adaptativo con lógica difusa. En este trabajo se utiliza la lógica difusa para ajustar las características del ruido del proceso (matrices de covarianza) que a priori se conocen de manera incompleta.
- “Study on FOG-SINS/GPS integrated attitude determination system using adaptive Kalman filter” (Chen, 2006) implementa la integración de GPS y sensor de movimiento (cabeceo y balanceo) para el monitoreo de la posición de una embarcación, mejorando la precisión de los datos de posición mediante filtrado adaptativo de Kalman. La adaptabilidad del filtro de Kalman se obtiene mediante el cálculo en tiempo real de las matrices de covarianza del sistema con ecuaciones recursivas.
- En “Fuzzy likelihood estimation based map matching for mobile robot self-localization”, (Jinxia, Zixing y Zhuohua, 2006) se combina estimación de máxima verosimilitud con lógica difusa para estimar la localización de un robot móvil en un ambiente 2D.
- “Optimal signal-adaptive maximum likelihood filter for enhancement of defects in eddy current C-scan images” (Hasanzadeh, Moghaddamjoo y Sadeghi, 2008) donde se requiere la detección de anomalías en materiales mediante imágenes obtenidas de la distribución superficial de la conductancia del material, requiriéndose la eliminación de ruido presente en toda la imagen. La relación conceptual con nuestro trabajo radica en la utilización de filtros para la estimación de valores constantes a partir de mediciones ruidosas, aunque en el trabajo citado utilizan una implementación recursiva de filtros adaptativos de máxima

verosimilitud tipo SAML⁵, MASML⁶ y Wiener para resolver el problema de estimación. El autor presenta resultados para los tres casos.

- En “Standard errors for EM estimation” (Jamshidian y Jennrich, 2000) se estudian variantes computacionales y sus correspondientes errores para la implementación del gradiente en el algoritmo de estimación de máxima verosimilitud EM.

Para el caso presente, la estrategia de estimación será discutida como parte de la propuesta sistémica de la investigación, a partir de las características presentadas por la información recolectada. De hecho, se encontró la necesidad de una solución al problema de la partición de un conjunto grande de datos en subgrupos ubicados en torno a un centroide, técnica conocida como partición espacial (spatial clustering).

En ese campo existe un amplio desarrollo teórico y práctico siendo el algoritmo *kmeans*, la técnica de mayor utilización, investigación y discusión. Por ejemplo, (Datta y Dougherty, 2007) lo establece junto al *kmeans* difuso (*Fuzzy Kmeans*), el *Self-organizing maps* y el particionamiento jerárquico (*Hierarchical Clustering*), como algoritmos estándar dentro del tema. Además, se encuentran trabajos con miras a superar sus limitaciones, como en el artículo “Fast and Robust General Purpose Clustering Algorithms” (Castro y Jang, 2004) donde se propone un algoritmo que mejora la estabilidad con respecto al ruido y valores errados (*outliers*) en comparación con el *kmeans* con un costo computacional levemente superior.

⁵ **SAML**: Signal Adaptative Filter base on Maximun Likelihood.

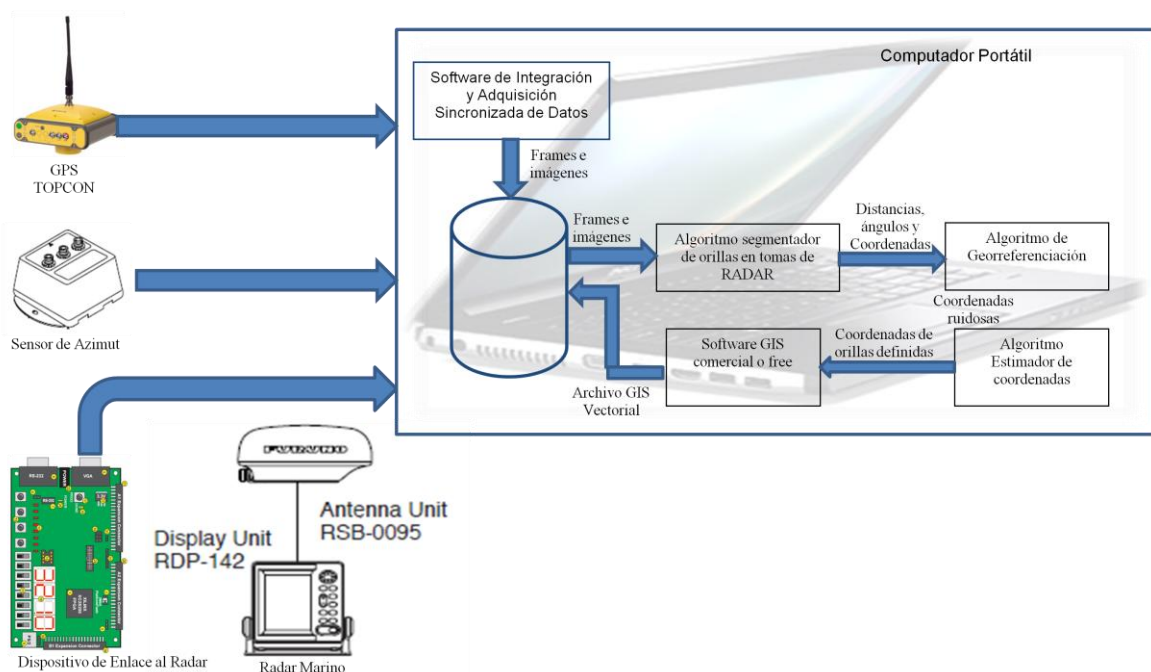
⁶ **MSAML**: Modified Adaptative Filter based on Maximun Likelihood.

3 IMPLEMENTACIÓN DEL SISTEMA DE GEORREFERENCIACIÓN

El proceso para la georreferenciación remota de orillas bajo la propuesta de esta tesis se desarrolla en dos etapas: La recolección (adquisición) de la información provista por los dispositivos que conformarán el sistema y el proceso de post-procesamiento en el que la información almacenada se retoma para aplicarle los algoritmos de georreferenciación y estimación que se propongan. A pesar de que la implementación práctica de la nueva técnica está basada en integración de tecnologías de amplio estado de desarrollo, parte de su componente investigativo se enmarca dentro de estudios exploratorio-descriptivos para descubrir algoritmos y configuraciones que permitan la integración de los componentes del sistema, de tal forma que sea viable la réplica del proceso y su explotación.

En esta sección se explican las implementaciones desarrolladas para la recolección y tratamiento de las fuentes de información. El esquema del sistema de recolección de información, georreferenciación y estimación de coordenadas se presenta en la Figura 3.1.

Figura 3.1. Diagrama de implementación del sistema de recolección de información y georreferenciación de esta tesis.



Las fuentes de información base se encontrarán a partir de las lecturas de distancias, azimut, coordenadas geográficas y tiempo entregada por los equipos electrónicos que hacen parte del sistema, los cuales son:

- Radar marino Furuno modelo M1712.
- FPGA Spartan III como tarjeta de desarrollo para el Dispositivo de Enlace al Radar.
- GPS Topcon Hyperlite plus
- Magnetic Heading Sensor (Sensor magnético de azimut) Furuno C500

La función desempeñada y la forma en que fueron construidos cada uno de los componentes del sistema se describen en las siguientes secciones.

3.1.1 Dispositivo de Enlace al Radar

Parte de la investigación de esta tesis consistió en determinar la estrategia adecuada para aprovechar la información de distancias y ángulos de detección de objetos que provee un radar comercial en su pantalla. El problema técnico surge debido a que el modelo de radar utilizado no provee en sus puertos de salida dicha información al no estar diseñado para este propósito. La anterior no es una limitación exclusiva de este radar, ya que en la consulta realizada al respecto no se encontró una marca o modelo que provea la información en la manera requerida

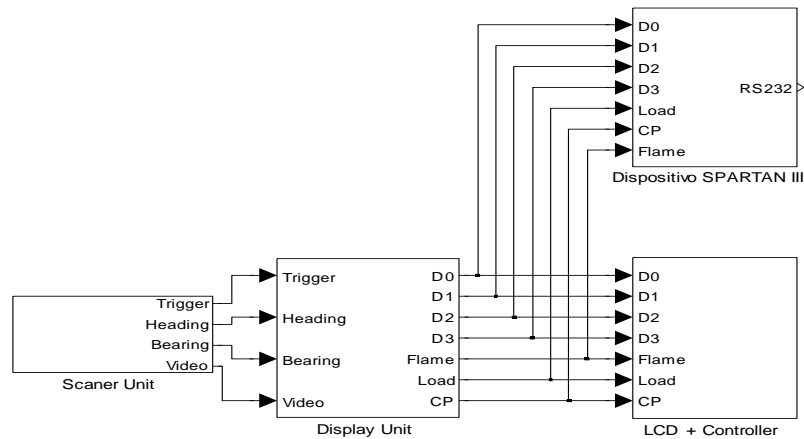
Por otro lado, es común que en los trabajos para extraer información adicional de un determinado radar se utilicen las señales bandabase de los pulsos emitidos (*Trigger*), los pulsos recibidos (*Video*), pulsos de sincronía de ángulo (*Heading*) y giro (*Bearing*) (Posada y Vélez, 2010), (Dankert y Rosenthal, 2004), (Serafino, 2010), debido a que en estas aplicaciones interesa información (como el *clutter*) que generalmente es eliminada por los bloques de procesamiento del equipo. En el caso presente, la opción más viable es aprovechar la información depurada por los sistemas de discriminación y umbralización de obstáculos realizado internamente por el radar debido a la complejidad y costo computacional que implican estos procesos. Se encontró más adecuado la utilización de la imagen de pantalla de radar, adicionando un dispositivo conectado en paralelo a las líneas de datos entre la unidad de procesamiento (*Display unit*) y la pantalla (*LCD+Controller*) como se ilustra en la Figura 3.2.

En este nodo la información de pantalla se provee mediante las señales (Furuno, 2000):

- *Flame*: pulso de sincronía que marca el inicio de datos para cada refresco de la pantalla.
- *Load*: pulso de sincronía que marca el inicio de cada una de las columnas de datos para el refresco de la pantalla.

- *CP*: pulso de sincronía cuyo flanco descendente corresponde con la disponibilidad de un nibble de bits para visualizar en pantalla.
- *D0-D3*: bus de cuatro líneas por donde se envían nibbles para su visualización en pantalla.

Figura 3.2. Diagrama de conexión del dispositivo de enlace al radar.

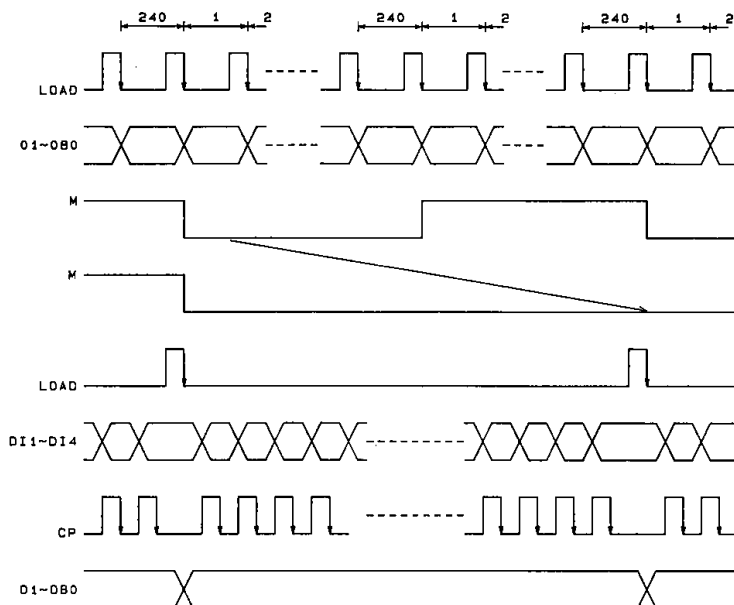


La temporización de estas señales se presenta en la Figura 3.3. El problema de adquisición para el radar se traduce bajo estas condiciones a obtener un dispositivo que utilizando las señales de pantalla del radar, las almacene y envíe hacia el sistema de georreferenciación. El procesador para dicho dispositivo debe tener la capacidad de realizar las tareas de almacenamiento de bits en sincronía con las señales de pantalla del radar, para mantener la integridad de la información visual y enviar dicha información.

Teniendo en cuenta las restricciones de tiempo en las tareas a realizar por este dispositivo que debe procesar 240x320 bits cada 13 ms a través de las cuatro líneas de datos a una tasa de 5Mbps cada una, se optó por la FPGA Spartan III que opera a 50MHz realizando una operación por ciclo, siendo este el factor determinante en la elección, además de la posibilidad de implementar un módulo de memoria RAM interna con dimensiones específicas para la aplicación disminuyendo el tiempo de diseño.

La configuración en lenguaje VHDL generó el dispositivo presentado en la Figura 3.4. De la figura se observa que este dispositivo está construido con bloques concurrentes y una máquina de estado encargada de controlar la ejecución del proceso encargado, cuyo diagrama de estados corresponde con la Figura 3.5.

Figura 3.3. Diagrama de tiempo de las señales de pantalla para el radar utilizado.



(Tomado de (Sanyo, 1995))

Figura 3.4. Diagrama de bloque equivalente al circuito interno del dispositivo de enlace al radar.

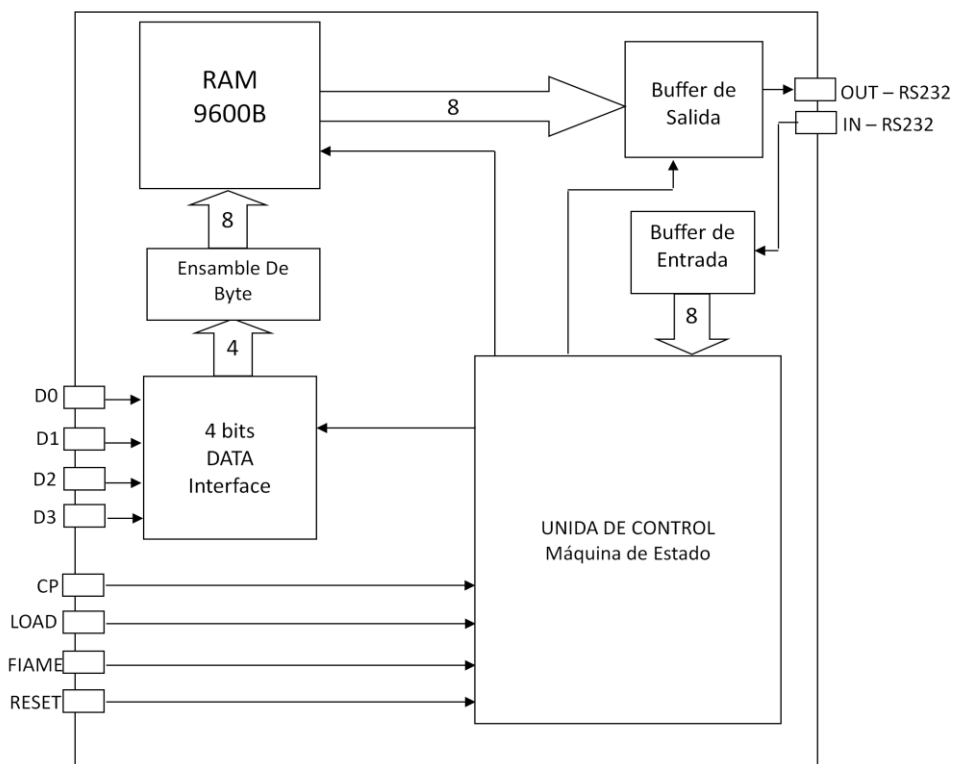
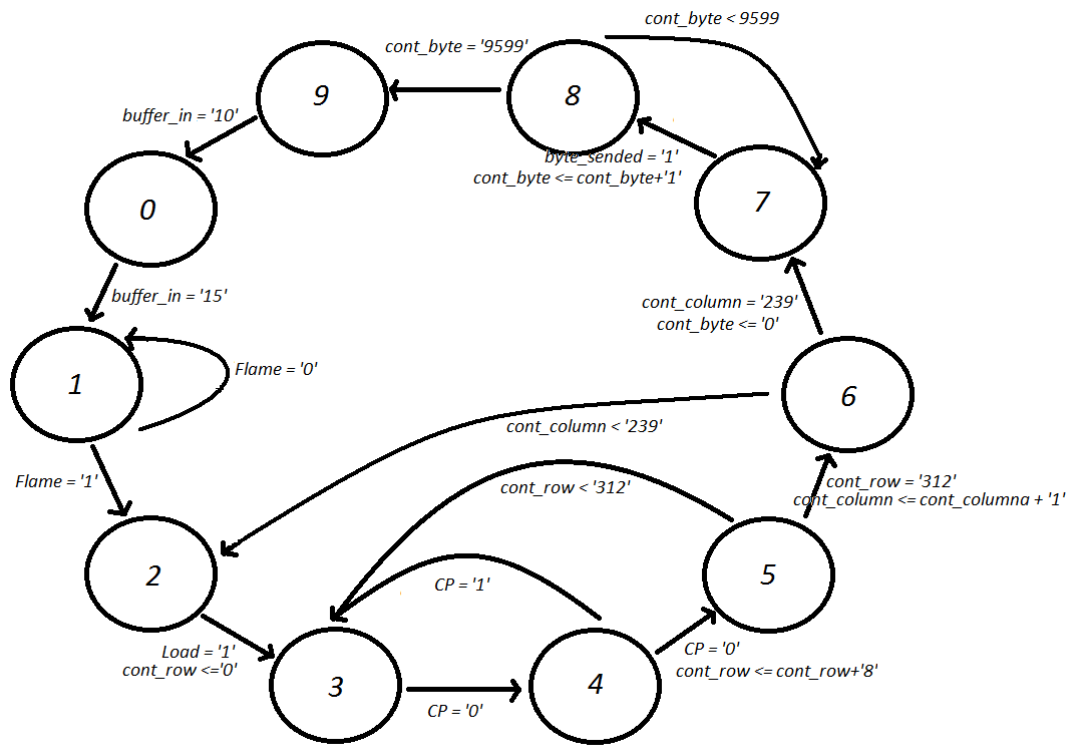


Figura 3.5. Diagrama de la Máquina de estados en Dispositivo de Enlace al Radar.



La secuencia de operación del dispositivo corresponde con el diagrama de estados de la Figura 3.5 que inicia en el estado “0” en espera de que se presente el byte ‘15’ en el búfer de entrada “*buffer_in*”, lo que permite el paso al estado “1” en el que permanece hasta que la señal “*Flame*” tome el valor ‘1’, que indica el inicio de transmisión de una imagen del radar a su pantalla. Al darse estas condiciones, el estado pasa a “2” donde permanece en espera del inicio de una columna a través de un ‘1’ en la señal “*Load*” y se mantiene en cero la variable contadora de filas “*cont_row*”. Al presentarse “*Load*” el estado del sistema pasa a “3” donde se aguarda hasta el flanco descendente de direccionar la entrada del bus de datos D0-D3 al nibble más y menos significativo del banco de memoria RAM alternativamente, además de incrementar el contador “*cont_row*”. Este proceso se realiza iterativamente entre los estados “3”, “4” y “5” hasta completar 320 bits de la columna, luego de lo cual se pasa al estado “6” donde se incrementa e inspecciona el contador de columnas “*cont_column*” de modo que se repita el proceso para una nueva columna desde el estado “2” o si están completadas 240, se transfiere el estado a “7” donde se inicia el envío de los bytes almacenados en RAM hacia el sistema de georreferenciación al irlos transfiriendo reiterativamente hacia el búfer de salida “*buffer_out*” del bloque UART para RS-323 configurado también en la FPGA. Al culminar el envío de 9600 bytes este sistema queda en el estado “9” a espera del byte “10” como confirmación para reiniciar el proceso desde el estado “0”.

3.1.2 Software de Integración y Adquisición Sincronizada de Datos

Dado que se utilizan como fuente de información tres equipos dispositivos independientes y asíncronos (Radar, GPS y sensor de azimut), es necesario la Adquisición y sincronización de la información de estos mediante una plataforma Hardware-Software que enlace las secuencias de datos de dichos instrumentos en estructuras de datos coherentes en tiempo y espacio. Dicha interfaz constituye un desarrollo específico para esta aplicación dado que debe realizar los siguientes procesos:

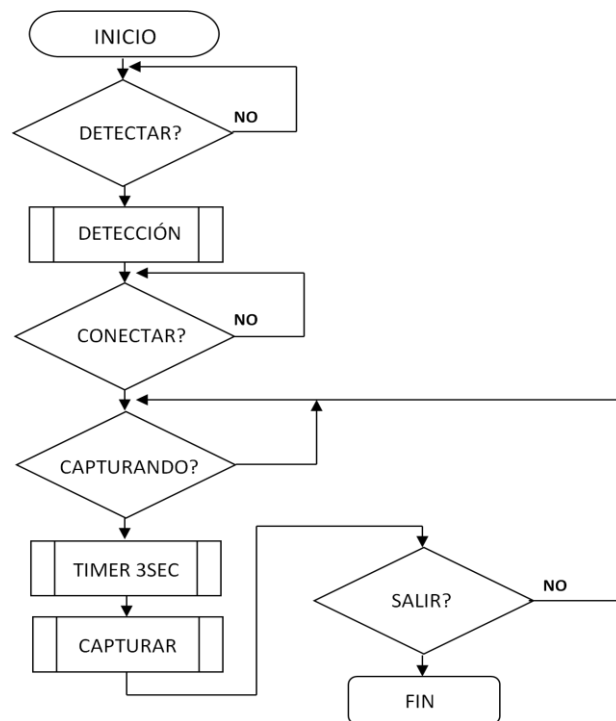
- Reconocimiento de los puertos seriales del computador, identificando que dispositivos del sistema están conectados en cada uno de estos.
- Manejo de los eventos de recepción para cada dispositivo, mediante programación por hilos se mantienen actualizadas las variables del GPS y el sensor de azimut.
- Proceso iterativo de captura y enlace de los datos entre dispositivos.
- decodificación de las secuencias de datos recibidas del radar en imágenes .png.
- decodificación de las sentencias NMEA entregadas por GPS y sensor de azimut, para extraer exclusivamente las variables de interés en esta aplicación: Longitud, Latitud, Altitud, tiempo de captura, velocidad, azimut real.
- Almacenamiento en disco duro de las tomas de radar, sincronizadas con líneas numeradas en las que se escriben las variables de posicionamiento. Las imágenes de radar se almacenan con un nombre que incluye el índice correspondiente al índice de su línea en un archivo de texto plano en el que se guardan los datos de GPS y sensor de azimut.

La ejecución de estos procesos se realiza en hilos de programación por eventos, propios del lenguaje en que se escribió la aplicación (C#). Sin embargo, por simplicidad se han descrito como una secuencia principal de operación presentada en el diagrama de flujo de la Figura 3.6.

Nótese que la secuencia ejecuta dos procedimientos principales: Detección y Captura que consisten en:

Detección: Antes de la conexión de los dispositivos el sistema, se ejecuta un proceso que escanea cada uno de los puertos seriales virtuales en existencia para el computador y ubica por cuál de estos se están transmitiendo sentencias NMEA, luego si en un determinado puerto se detectan las sentencias, se procede a su análisis para establecer el tipo de dispositivo al que está accediendo. Si la conexión es exitosa bajo los parámetros y puertos seleccionados, dicha configuración se almacena en una variable o tabla de conexiones, de modo que este proceso no deba repetirse. El diagrama de flujo correspondiente se presenta en la Figura 3.7a).

Figura 3.6. Diagrama de flujo principal del software de adquisición.



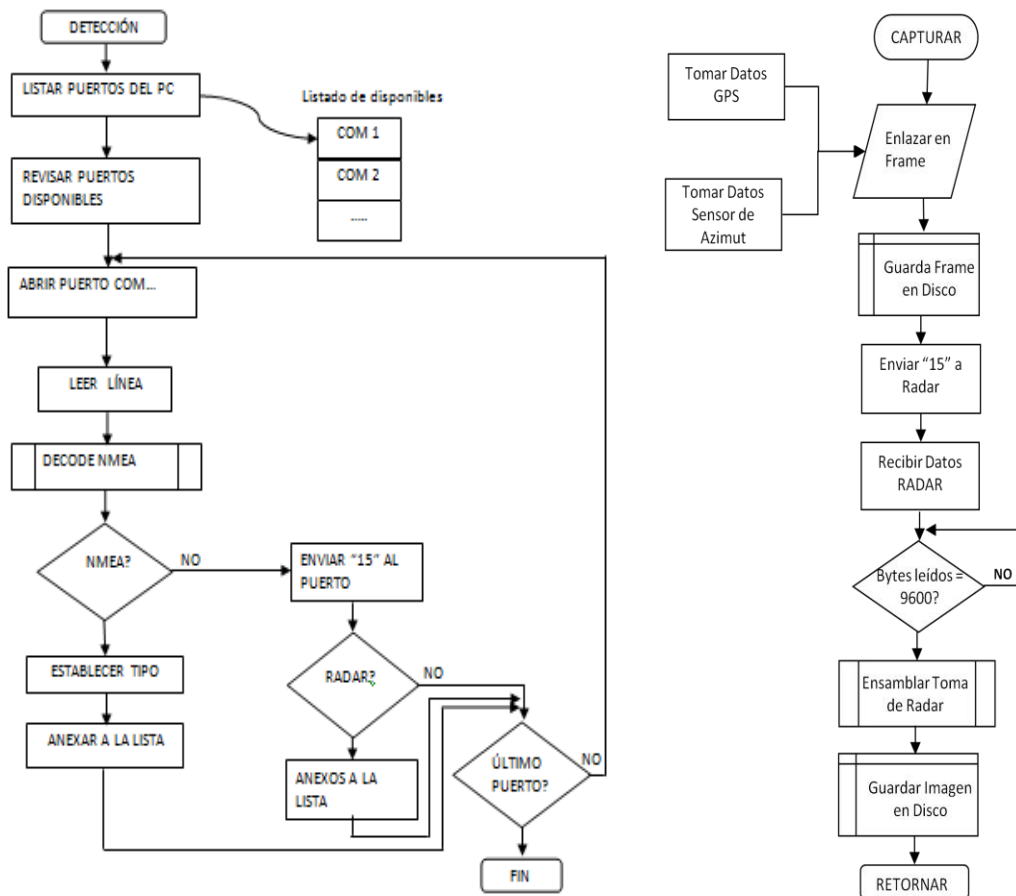
Captura: cada vez que se invoca a este procedimiento, se ejecuta de acuerdo a la secuencia presentada en el diagrama de flujo de la Figura 3.7b), donde la toma de datos consiste en el uso de las variables de posicionamiento que han sido almacenadas en memoria RAM por los procesos asíncronos de *Recepción y decodificación NMEA* para el GPS y azimuth, de los cuales se maneja un índice que se incrementa si llega una nueva coordenada. Este índice sirve de guía para toda la información del proceso, ya que se utiliza como marcador de las filas o frames que es la forma como se almacena esta información (Véase Tabla 3.1) antes de guardarla en el disco duro como archivo de texto con tabulaciones.

Posteriormente, se solicita al *Dispositivo de Enlace al Radar* (Véase Sec. 3.3.2) el envío de la información de radar mediante el Byte “15”, abriendo el puerto correspondiente y se ejecuta la instrucción de lectura. Este dispositivo es el único al que se le capturan los datos de forma síncrona, debido a que se debe controlar la consistencia del frame de datos recibidos, el cual se pasará por el proceso de *Ensamblar toma de Radar*. Finalmente se envía al dispositivo de acople al radar el byte “10” para que quede en espera de una nueva toma, que depende de que se repita todo el proceso de *Captura* en un nuevo evento generado por el “timer” de repetición cada 3 segundos, teniendo en cuenta que el radar genera una nueva toma con una periodicidad cercana a este valor.

Tabla 3.1. Organización de datos de GPS y Azimut Almacenados

Índice	Longitud	Latitud	Altitud	Hora	Azimut	Velocidad
1	-74,8407451	11,0641997	5,5941	20:16:26	102,322	0,5391
2	-74,840743	11,0641958	5,5606	20:16:29	207,318	0,611
4	-74,8407552	11,0641795	5,5391	20:16:35	173,865	2,8057
5	-74,8407499	11,0641479	5,5144	20:16:38	167,374	5,3772
6	-74,8407339	11,0641052	5,4842	20:16:41	150,796	6,4923

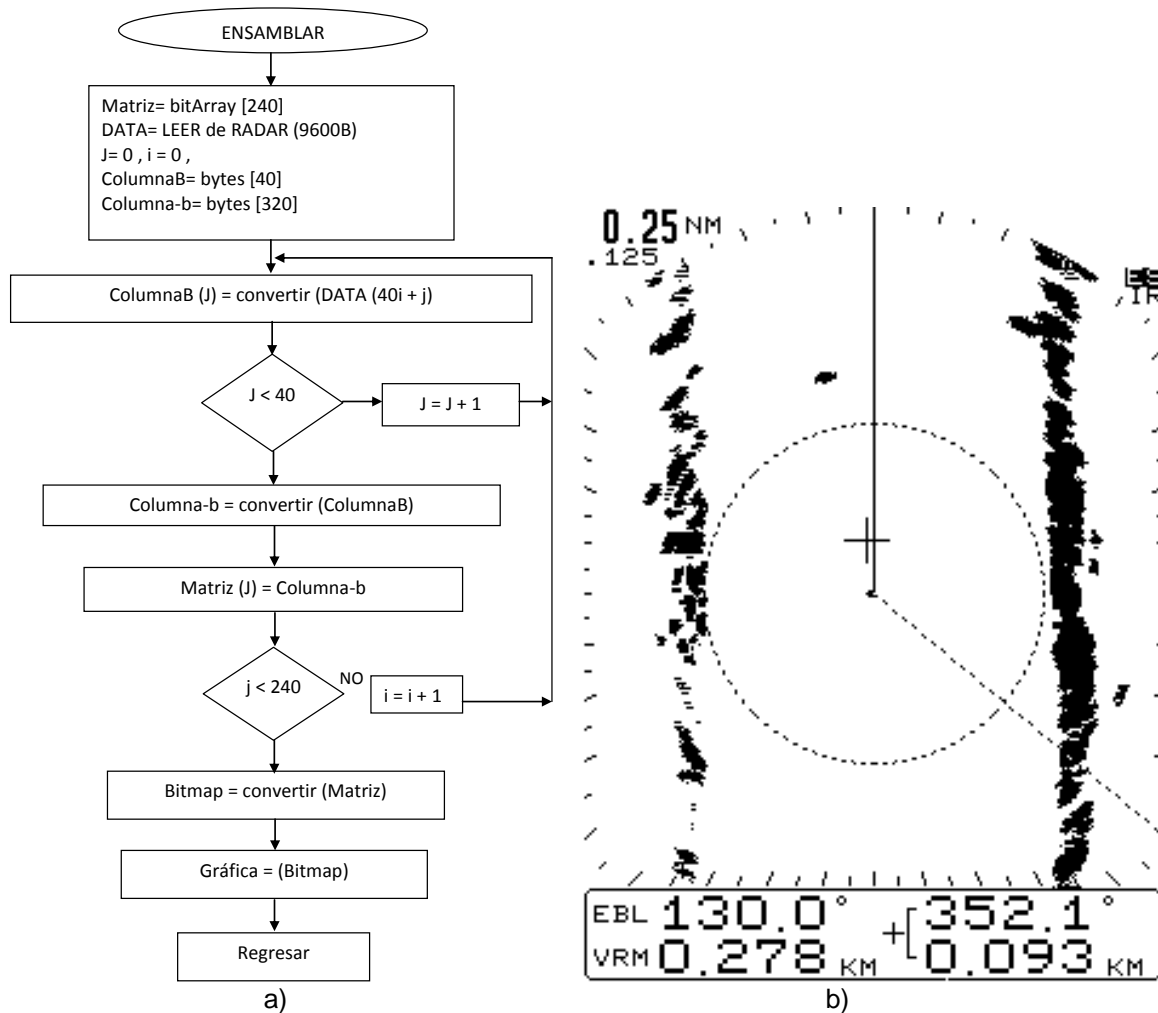
Figura 3.7 .a) Diagrama de flujo del proceso de captura. **b)** Diagrama de flujo del proceso de ensamble de imágenes de radar.



Ensamblar Toma de Radar: Los datos recibidos vía puerto serial RS-232 no representan en sí una imagen, debido a que no tiene las dimensiones ni las características propias de estas, lo que se puede corroborar al revisar el proceso realizado en el *Dispositivo de Enlace al Radar*. En este proceso una vez se han recibido los 9600 Bytes de la información de radar, se segmentan en grupos de 40 Bytes correspondientes a la información de una columna de la pantalla, las cuales se pueden ir convirtiendo en

vectores de bits (*BitArrays*) y agrupando en una matriz con 240 de estos elementos, lo que equivale a tener una matriz de 320x240 bits similar a la información de pantalla del radar. Esta matriz de bits se utiliza como fuente de información para asignar el color blanco o negro en los distintos pixeles de un mapa de bits, el cual finalmente es proyectado en la interfaz gráfica del software de integración y almacenada bajo la codificación .png. El diagrama de flujo del proceso que se acaba de describir, es presentado en la Figura 3.8a), generando imágenes como la presentada en la Figura 3.8b).

Figura 3.8. a) Diagrama de flujo del proceso Ensamblar Toma de Radar. b) Imágenes de Radar generadas



3.1.3 Algoritmo de Georreferenciación

Un componente importante del sistema será el desarrollo del software en el que se apliquen los procedimientos geométricos para georreferenciación y los algoritmos de reducción de incertidumbre y diezmado de los datos de entrada. La solución a este problema será desarrollada a nivel teórico en el capítulo 4. El algoritmo debe ir tomando cada frame de datos GPS y Azimut, cargar la correspondiente toma de radar, aplicar la medición de distancias y ángulos sobre la imagen de radar, y a partir de los datos sincronizados calcular las coordenadas referentes a cada distancia y ángulo registrado.

Dichas coordenadas conformarán un conjunto ruidoso de muestras, cuyo procesamiento estadístico permitirá la obtención de un grupo reducido de coordenadas pero que describen satisfactoriamente la morfología de la orilla.

4 DESCRIPCIÓN TEÓRICA DEL SISTEMA PROPUESTO

En el capítulo 1 se definió el problema científico que tratamos de resolver en esta tesis, delimitado con los objetivos específicos planteados. En el presente capítulo se expone la propuesta de solución al problema de investigación de la tesis. En la primera sección se expone el modelo matemático obtenido del sistema de georreferenciación. Sin embargo, dada la incertidumbre propia de las fuentes de información de entrada, la naturaleza del proceso es aleatoria con alta cantidad de muestras redundantes, por lo que en la sección posterior se explican las técnicas de procesamiento estadístico utilizadas para la reducción y aprovechamiento de la redundancia en pro de mejorar la exactitud de las coordenadas de orilla obtenidas. Finalmente se presentan las métricas de desempeño utilizadas para calificar la solución obtenida.

4.1 MODELO MATEMÁTICO PARA LAS COORDENADAS DE SALIDA

La propuesta principal de este trabajo consiste en la georreferenciación remota de las orillas fluviales, al tomar los distintos puntos de la orilla como los objetos a detectar por un radar marino desde una embarcación. Para conseguirlo es necesario plantear un sistema de conversión matemática que calcule las coordenadas geográficas de objetos a distancia, transformando la referencia espacial distancia-ángulo suministrada por el radar en una referencia geocéntrica.

La Figura 4.1 presenta un esquema que ejemplifica, de manera simplificada, la situación típica para obtener la coordenada geográfica de un punto en la orilla a partir de la información que entregarían el radar, GPS y sensor de azimuth o brújula electrónica. De la figura, θ es el ángulo de azimuth dado por el sensor correspondiente, R la distancia a la orilla medida por un radar o ecosonda y $(\lambda_{GPS}, \phi_{GPS})$ son Longitud y Latitud de la coordenada geográfica de la embarcación en el estándar WGS84. Por el carácter bidimensional del gráfico, no se ha representado la Altura elipsoidal h_{GPS} que hace parte de la coordenada geográfica.

En la Figura 4.2, se presenta la proyección de la posición y orientación de la embarcación, describiendo un sistema de referencia cartesiano que se ha corregido en un ángulo θ para quedar orientado al norte geográfico. Esta contextualización geográfica se conoce como proyección cartesiana (IGAC, 2004).

Figura 4.1. Situación geométrica simplificada para obtención de coordenadas geográficas de un punto en la orilla.

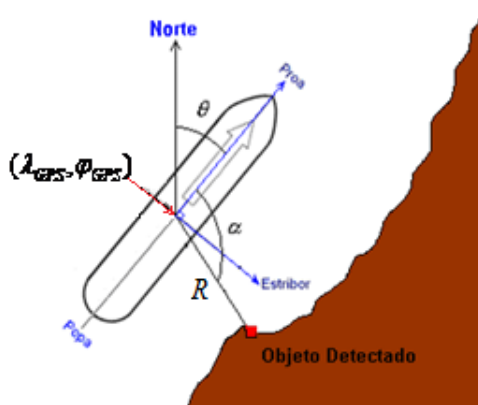
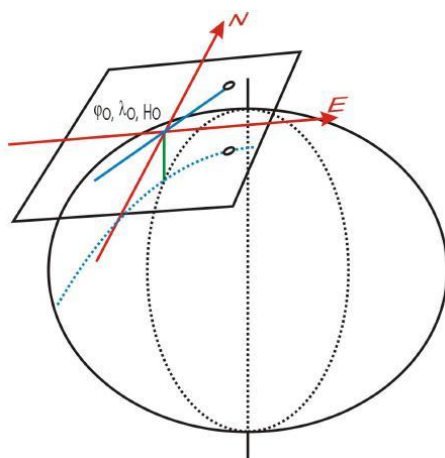


Figura 4.2. Modelo de proyección plana cartesiana sobre elipsoide terrestre.



Tomado de (IGAC, 2004).

Bajo ese sistema rectangular local de referencia, un objeto detectado (orilla, barco, isla, piedras, etc) se posiciona en la dirección oeste mediante la expresión (4.1.1) y en la dirección norte mediante (4.1.2).

$$d_o = R \cdot \text{Sen}(\theta + \alpha) \quad (4.1.1)$$

$$d_N = R \cdot \text{Cos}(\theta + \alpha) \quad (4.1.2)$$

El proceso de georreferenciación, implica llevar esta referencia local a un marco geocéntrico, con el fin de enmarcar esta ubicación en un sistema de referencia estándar como el WGS84. De primera impresión, este problema parecería estar resuelto en (IGAC, 2004) mediante las transformaciones afines desde coordenadas planas a elipsoidales

para los distintos datum del país. En el caso de la presente aplicación, la movilidad de la embarcación hace inviable la aplicación de dichos modelos de conversión, los cuales tienen riesgo de imprecisiones y desviaciones sistemáticas si no se configuran adecuadamente. Además, la proliferación de Datums y modelos de proyección plana incluso para una misma zona, hacen que incluso la configuración de los softwares de conversión sea una labor rigurosa y tediosa.

A continuación se desarrollará una solución al problema de georreferenciación desde coordenadas o referencias en sistemas de proyección cartesiana, encontrando una función de transformación que varía con la ubicación de estos sobre cualquier lugar del elipsoide terrestre. El primer supuesto del modelo consiste en asumir que las orillas están ubicadas sobre el mismo plano de la embarcación, el cual es tangente al elipsoide terrestre siguiendo la convención de la proyección cartesiana. A partir de este supuesto, las coordenadas geocéntricas cartesianas x , y , z de un punto en la orilla son tales que cumplen los criterios:

$$n_x(x - x_{GPS}) + n_y(y - y_{GPS}) + n_z(z - z_{GPS}) = 0 \quad (4.1.3)$$

$$\sqrt{(x - x_{GPS})^2 + (y - y_{GPS})^2 + (z - z_{GPS})^2} = D^2 \quad (4.1.4)$$

$$T_x(x - x_{GPS}) + T_y(y - y_{GPS}) + T_z(z - z_{GPS}) = D|\vec{T}|\cos(\theta + \alpha) \quad (4.1.5)$$

Siendo (4.1.3) la ecuación del plano tangente al elipsoide terrestre, con vector normal dado por las componentes $\{n_x, n_y, n_z\}$, (4.1.4) la relación de distancia euclidiana entre las coordenadas a calcular y el origen del sistema de proyección cartesiana dado por las coordenadas rectangulares del GPS (x_{GPS} , y_{GPS} , z_{GPS}) y (4.1.5) expresa la condición de dirección (producto punto) entre el vector tangente al plano que apunta al Norte $\{T_x, T_y, T_z\}$ y la dirección con que fue avistada el punto en la orilla.

Las coordenadas rectangulares del GPS son obtenidas al evaluar la ecuación de transformación (2.3.1) en la longitud λ_{GPS} , la latitud φ_{GPS} y la altura elipsoidal h_{GPS} entregadas por este dispositivo (véase (4.1.6) y (4.1.7)).

$$\begin{bmatrix} x_{GPS} \\ y_{GPS} \\ z_{GPS} \end{bmatrix} = \begin{bmatrix} (\eta_{GPS} + h_{GPS}) \cos \varphi_{GPS} \cos \lambda_{GPS} \\ (\eta_{GPS} + h_{GPS}) \cos \varphi_{GPS} \sin \lambda_{GPS} \\ ((1 - e^2) \eta_{GPS} + h_{GPS}) \sin \varphi_{GPS} \end{bmatrix} \quad (4.1.6)$$

$$\eta_{GPS} = \frac{a}{\sqrt{1 - e^2 \sin^2 \varphi_{GPS}}} \quad (4.1.7)$$

Además (4.1.3), (4.1.4) y (4.1.5) expresan un sistema de ecuaciones en términos de los desplazamientos en el espacio tridimensional entre el origen del sistema de proyección cartesiana dada y la coordenada buscada:

$$\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} x - x_{GPS} \\ y - y_{GPS} \\ z - z_{GPS} \end{bmatrix} \quad (4.1.8)$$

De modo que el problema ahora consiste en encontrar los vectores unitarios en el espacio tridimensional que sirvan de dirección a los desplazamientos d_{XN} y d_{YN} para transformarlos en $(\Delta x, \Delta y, \Delta z)$ adecuados para que (x, y, z) cumpla los criterios dados por (4.1.3), (4.1.4) y (4.1.5).

Para ello, permitamos expresar a (2.2.1) como una función que genera puntos x, y, z en toda la superficie del elipsoide terrestre:

$$\begin{aligned} \vec{r}(\lambda, \varphi, h) &= x\hat{i} + y\hat{j} + z\hat{k} = \\ &(\eta(\varphi) + h)\cos\varphi\cos\lambda\hat{i} + (\eta(\varphi) + h)\cos\varphi\sin\lambda\hat{j} + ((1 - e^2)\eta(\varphi) + h)\sin\varphi\hat{k} \end{aligned} \quad (4.1.9)$$

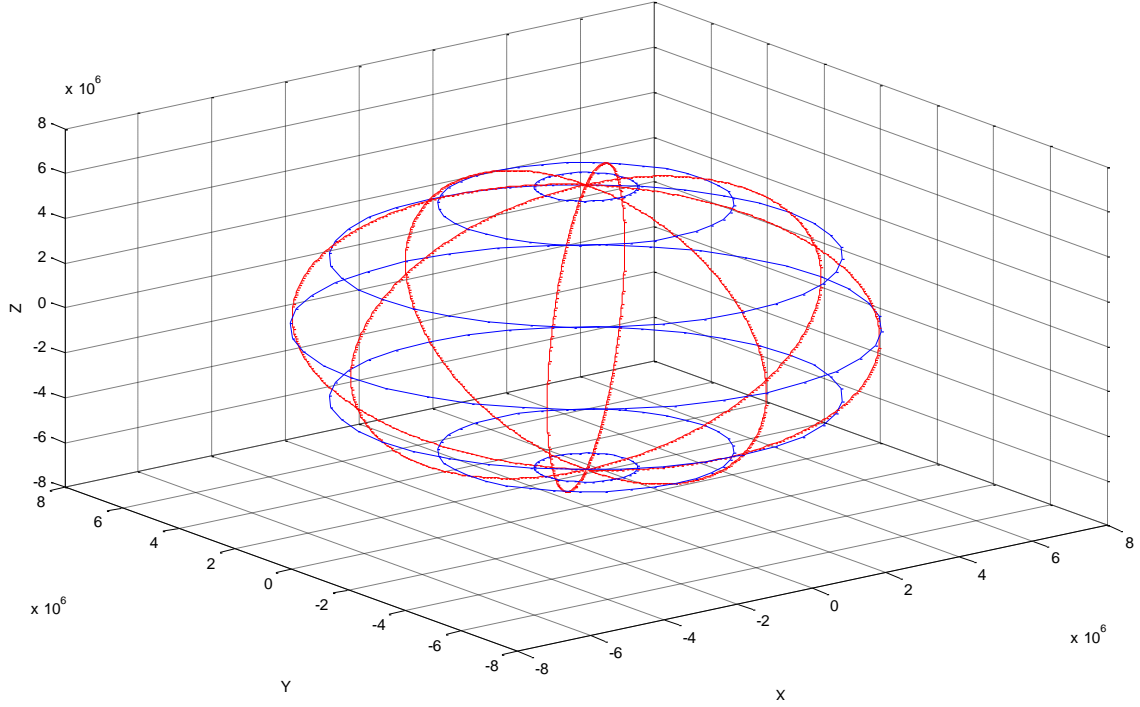
Nótese que (4.1.9) generaliza a un conjunto de curvas paramétricas en los intervalos $-90 < \varphi < 90$ y $-180 < \lambda < 180$. Donde $\vec{r}(\lambda)|_{(\varphi_{GPS}, h_{GPS})}$ y $\vec{r}(\varphi)|_{(\lambda_{GPS}, h_{GPS})}$ serían las curvas del paralelo y meridiano que pasan por $(\lambda_{GPS}, \varphi_{GPS}, h_{GPS})$ y cumplen la condición de ser tangentes al elipsoide terrestre aportando direcciones hacia el Norte y el Oeste buscadas en el espacio tridimensional geocéntrico. La Figura 4.3 ilustra la generación de algunos meridianos y paralelos sobre el elipsoide terrestre variando los parámetros a (4.1.9).

Para una curva paramétrica los vectores tangentes en un determinado punto se definen evaluando la primera derivada con respecto al parámetro en dicho punto (Steward, 1995). En este caso, aplicando dicho criterio, los vectores unitarios hacia el norte y oeste se obtienen respectivamente como:

$$\hat{N}(\lambda, \varphi, h) = \frac{\partial \vec{r} / \partial \varphi}{\left| \partial \vec{r} / \partial \varphi \right|} \quad (4.1.10)$$

$$\bar{O}(\lambda, \varphi, h) = \frac{\partial \vec{r} / \partial \lambda}{\left| \partial \vec{r} / \partial \lambda \right|} \quad (4.1.11)$$

Figura 4.3. Elipsoide terrestre generado a partir de curvas paramétricas evaluadas en $\lambda = [0, 45, 90, 135, 180, 225, 270, 315]$ y $\varphi = [-80, -60, -30, 0, 30, 60, 80]$.



Luego, las derivadas en (4.1.10) y (4.1.11) se resuelven obteniéndose:

$$\begin{bmatrix} \frac{\partial \vec{r}}{\partial \varphi} \hat{i} \\ \frac{\partial \vec{r}}{\partial \varphi} \hat{j} \\ \frac{\partial \vec{r}}{\partial \varphi} \hat{k} \end{bmatrix} = \begin{bmatrix} \cos(\lambda) \left[\frac{a \cdot e^2 \cdot \sin(\varphi) \cos^2(\varphi)}{(1 - e^2 \cdot \sin^2(\varphi))^{3/2}} - \frac{a \cdot \sin(\varphi)}{(1 - e^2 \cdot \sin^2(\varphi))^{1/2}} - h \cdot \sin(\varphi) \right] \hat{i} \\ \sin(\lambda) \left[\frac{a \cdot e^2 \cdot \sin(\varphi) \cos^2(\varphi)}{(1 - e^2 \cdot \sin^2(\varphi))^{3/2}} - \frac{a \cdot \sin(\varphi)}{(1 - e^2 \cdot \sin^2(\varphi))^{1/2}} - h \cdot \sin(\varphi) \right] \hat{j} \\ \cos(\varphi) \left[h - \frac{a \cdot (1 - e^2)}{(1 - e^2 \cdot \sin^2(\varphi))^{3/2}} \right] \hat{k} \end{bmatrix} \quad (4.1.12)$$

$$\begin{bmatrix} \frac{\partial \vec{r}}{\partial \lambda} \hat{i} \\ \frac{\partial \vec{r}}{\partial \lambda} \hat{j} \\ \frac{\partial \vec{r}}{\partial \lambda} \hat{k} \end{bmatrix} = \begin{bmatrix} -\sin(\lambda) \cdot \cos(\varphi) \left[\frac{a}{(1 - e^2 \cdot \sin^2(\varphi))^{1/2}} + h \right] \hat{i} \\ \cos(\lambda) \cdot \cos(\varphi) \left[\frac{a}{(1 - e^2 \cdot \sin^2(\varphi))^{1/2}} + h \right] \hat{j} \\ 0 \end{bmatrix} \quad (4.1.13)$$

Al disponer de los vectores de dirección al Norte y Oeste sobre el plano de proyección cartesiana, las coordenadas de un punto en la orilla (x_p, y_p, z_p) se obtendrán como:

$$\begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} = \begin{bmatrix} x_{GPS} \\ y_{GPS} \\ z_{GPS} \end{bmatrix} + d_N \cdot \vec{N}(\lambda_{GPS}, \varphi_{GPS}, h_{GPS}) + d_O \cdot \vec{O}(\lambda_{GPS}, \varphi_{GPS}, h_{GPS}) \quad (4.1.14)$$

La ecuación (4.1.14) constituye un modelo matemático para relacionar las entradas λ_{GPS} , φ_{GPS} , h_{GPS} , R , θ y α con la salida (x_p, y_p, z_p) , evaluando previamente a las ecuaciones (4.1.1), (4.1.2), (4.1.6), (4.1.7), (4.1.10) y (4.1.11) en los valores correspondientes. Finalmente, las coordenadas en el formato WGS84 se obtendrían aplicando la conversión de rectangulares a elipsoidal:

$$\begin{bmatrix} \lambda_p \\ \varphi_p \\ h_p \end{bmatrix} = \begin{bmatrix} \arctan\left(\frac{y_p}{x_p}\right) \\ \arctan\left(\frac{z_p + e^2 b \sin^3 \vartheta}{\sqrt{x_p^2 + y_p^2} - e^2 a \cos^3 \vartheta}\right) \\ \frac{\sqrt{x_p^2 + y_p^2}}{\cos \varphi} - \frac{a}{\sqrt{1 - e^2 \sin^2 \varphi}} \end{bmatrix} \quad (4.1.15)$$

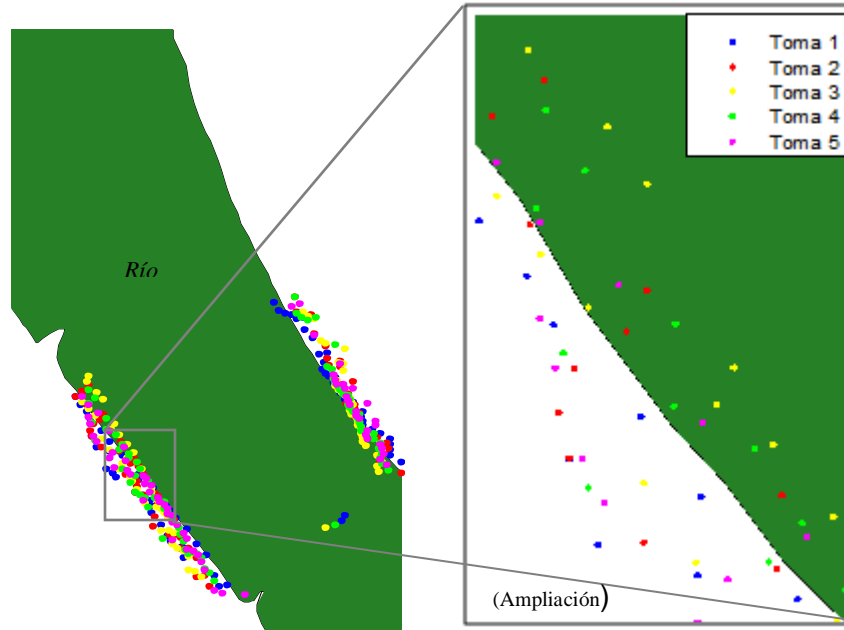
$$\vartheta = \arctan\left(\frac{z_p a}{\sqrt{x_p^2 + y_p^2} b}\right) \quad (4.1.16)$$

4.2 ESTIMACIÓN DE COORDENADAS DE LA ORILLA

En la sección 5.1.3 se mostrará cómo el algoritmo de georreferenciación produce una alta cantidad de muestras desviadas alrededor de la coordenada real de un punto detectado. La ocurrencia de este fenómeno es consecuencia natural de las múltiples detecciones que el radar hace de un mismo sector de la orilla. Al superponerse las coordenadas generadas por cada toma, la orilla se presenta como objeto difuso en el que se entremezclan distintas estimaciones de sus puntos.

En la Figura 4.4 se muestran los puntos de orilla generados por cinco (5) tomas del radar desde las cinco (5) primeras posiciones de la embarcación en movimiento. De la Figura citada, nótese que los puntos de orilla generados en cada toma se identifican con un color distinto, mostrando que finalmente la información se entremezcla de manera impredecible.

Figura 4.4. Puntos o coordenadas de la orilla generadas en primeras tomas de radar.



La anterior situación se presenta a pesar de que el modelo matemático de georeferenciación es completamente determinístico, pero debe recordarse que las entradas (GPS, radar, sensor de azimuth) son instrumentos electrónicos cuyos datos son en realidad estimaciones de la variable que pretenden medir. Además, revisando la situación en que se presentan las mediciones, se encontrará que un mismo punto es registrado varias veces por el radar con distintos valores de distancia y ángulo a medida que avanza la embarcación, cada una de estas mediciones con alta incertidumbre.

De esta manera, surge el problema de encontrar la estimación de cada una de las coordenadas en la orilla, disminuyendo la cantidad de muestras a las necesarias para garantizar una resolución determinada, a partir del aprovechamiento de la redundancia de la información alrededor de un punto detectado. Formalmente, sea S el conjunto de todas las coordenadas del proceso:

$$S = \{\vec{s}_1, \vec{s}_2, \dots, \vec{s}_n\} \quad (4.2.1)$$

Siendo n el total de coordenadas obtenidas y cada $s_i = (\lambda_p(i), \varphi_p(i), h_p(i))$. De este conjunto se deben extraer K grupos de tal forma que se formen los conjuntos:

$$X_k = \{\vec{s}_1(k), \vec{s}_2(k), \dots, \vec{s}_{l_k}(k)\} \subset S \quad (4.2.2)$$

Donde l_k corresponde a la cantidad de coordenadas dispuestas en el k -ésimo grupo. Note que en todo caso:

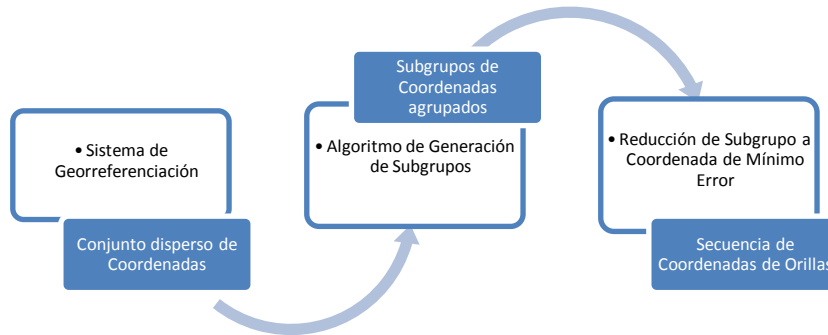
$$S = X_1 \cup X_2 \cup X_3 \cup \dots X_K \quad (4.2.3)$$

Luego el conjunto de la estimación de las coordenadas de la orilla se obtendría como:

$$\hat{O} = \{\Psi(X_1), \Psi(X_2), \Psi(X_3) \dots \Psi(X_K)\} \quad (4.2.4)$$

Siendo Ψ una función estimadora de cada una las coordenadas finales $(\hat{\lambda}_p(k), \hat{\phi}_p(k), \hat{h}_p(k))$ a partir de las l_k muestras de ese subgrupo. De esta manera, el proceso de estimación propuesto para el sistema de georreferenciación está compuesto de dos etapas: la determinación de los X_K subgrupos y la aplicación de función estimadora Ψ sobre cada uno de estos (véase Figura 4.5).

Figura 4.5. Integración de Sistema de Georreferenciación con estimación de coordenadas finales de la orilla.



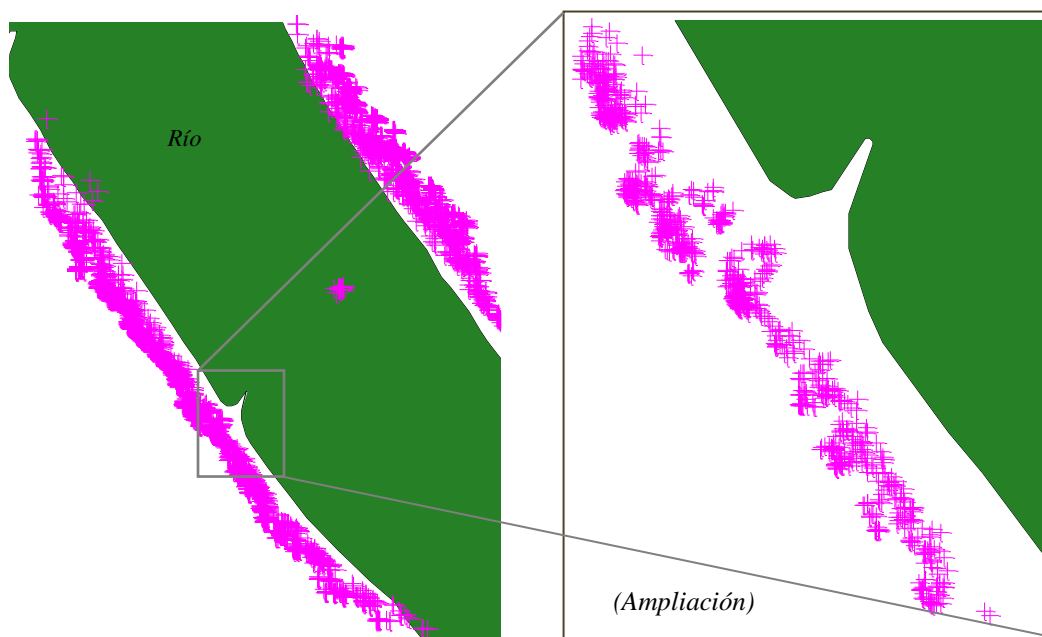
La solución para que la secuencia de coordenadas obtenidas sea una estimación de buena calidad de las coordenadas exactas de la orilla se resolverá en las secciones 4.2.1 y 4.2.2.

4.2.1 Generación de Subgrupos (Cluster's) para Puntos Significativos de la Orilla

Una propuesta de solución a este problema fue planteado en (Jiménez y Posada, 2007), donde se proponía tomar cada una de las coordenadas y encontrar el conjunto de coordenadas “hermanas” es decir aquellas que posiblemente corresponden al mismo punto exacto de la orilla y que fueron generadas por la detección del radar desde múltiples posiciones. La implementación de esta solución requiere el cálculo inverso al proceso de georreferenciación, es decir determinar desde cada una de las coordenadas del proceso, con cuál ángulo y distancia fue detectado ese punto con el radar y en

consecuencia seleccionar las otras coordenadas del proceso que tenga mayor correspondencia con los ángulos y distancias estimados (Véase Anexo 1). Dicha propuesta fue implementada computacionalmente y aplicada a las muestras generadas por nuestro proceso, obteniéndose resultados de la forma mostrada en la Figura 4.6. En la figura, las marcas fucsia representan la solución generada por el algoritmo citado.

Figura 4.6. Puntos o coordenadas de orilla generadas en primera toma de radar.



Al revisar estos resultados, se observa que la propuesta del anexo no soluciona satisfactoriamente el problema, lo cual se debe a que en dicha iniciativa se preveía una detección de radar más idealizada (modelo de rayos), mientras que los datos generados por el proceso real develan una interacción más compleja. Por ejemplo, al considerar el efecto del ancho del patrón de radiación a las distancias implicadas, para la distancia máxima registrada (456 m), el ancho de la superficie de incertidumbre es de 40 m, lo que implica que cada punto de las muestras realmente representa un área importante a su alrededor.

En consecuencia, se enfocará el problema de determinación de los grupos asumiendo un carácter estocástico en su procedencia, es decir se tomará cada una de las coordenadas como muestra o indicador de cercanía a un objeto de nuestro interés con posición desconocida: la orilla. Bajo esta premisa, se está ignorando, de momento, cómo pudo haber llegado cada coordenada del proceso allí, estando asociadas solamente en función de su ubicación espacial.

Para este tipo de problema, se hace necesaria la utilización de alguna técnica de partición espacial, debido al volumen y disgregación presentada por las coordenadas. La técnica *Kmeans* es una de las más aceptada en la literatura para resolver el problema de partición espacial de grupos (Draghici, 2003), (Mirkin, 2005) en volúmenes grandes de datos. Este algoritmo, es una aproximación de la solución óptima para el problema de optimización:

$$C = \arg_C \min \left| \sum_{i=1}^n Dist(\vec{s}_i, Closest(C, \vec{s}_i)) \right| \quad (4.2.5)$$

Donde

- $C = \{\vec{c}_1, \vec{c}_2, \dots, \vec{c}_K\}$ Es el conjunto de “centroides” alrededor de los cuales se establecerán los subgrupos X_k del conjunto principal S .
- $Closest(C, \vec{s}_i)$ Es la operación que selecciona el elemento \vec{c}_k más cercano a \vec{s}_i .
- Y $Dist(\vec{s}_i, \vec{c}_k)$ define el operador de distancia espacial utilizado.

Aunque hay autores que limitan el *kmeans* a la optimización de distancia euclidiana (Datta y Dougherty, 2007), es posible seleccionar otros tipos de distancia espacial, modulando el comportamiento del algoritmo (Draghici, 2003). Mediante experimentación, se determinó que los tipos de distancias a utilizar en este trabajo deben ser:

- *Distancia Euclidiana*: sea \vec{x} y \vec{y} vectores en el espacio de r dimensiones, la distancia euclidiana o pitagórica se calcula como:

$$d_E(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^r (x_i - y_i)^2} \quad (4.2.6)$$

- *Distancia Manhattan o “city”*: calculada como la suma de las diferencias absolutas, de tal forma que para los vectores del caso anterior:

$$d_M(\vec{x}, \vec{y}) = \sum_{i=1}^r |x_i - y_i| \quad (4.2.7)$$

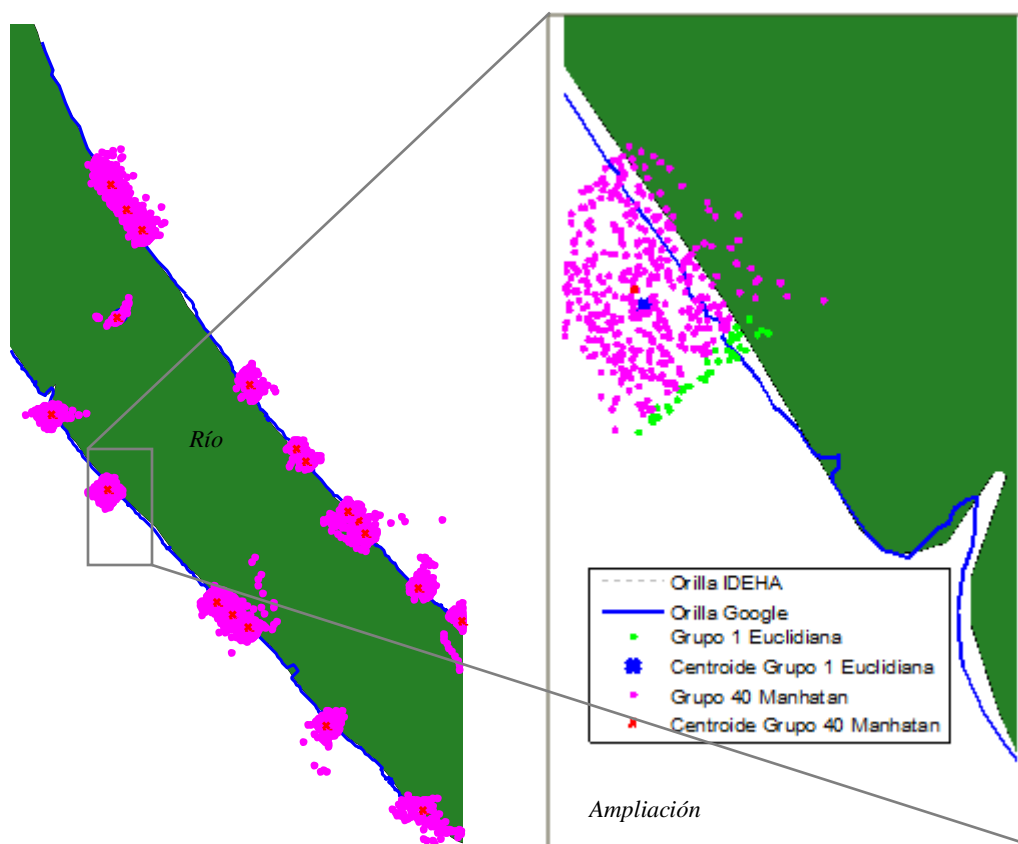
Para el algoritmo de *kmeans*, una característica importante son los problemas de aleatoriedad en los resultados de salida debido a su dependencia en la semilla C_I o solución inicial para la inicialización de las iteraciones del algoritmo. Actualmente, no ha sido posible vislumbrar una solución inicial válida del proceso, más aún si se tiene en cuenta que en nuestro contexto, habrá lugares del río con orillas que deben asumirse en absoluto desconocidas, debido a la aparición de islas y los cambios bruscos generados por inundaciones y sequías. Teniendo en cuenta lo anterior, la semilla inicial para el

algoritmo *kmeans* será aleatoria dejando la evaluación de las soluciones en el parámetro K de la cantidad de grupos y en los tipos de distancias definidas. Los resultados de la experimentación con variantes en la cantidad de subgrupos y el tipo de distancia en el algoritmo *kmeans* se detallan en la sección 5.2.

4.2.2 Reducción de Subgrupos a Coordenadas de menor Incertidumbre

En la Figura 4.7 se presenta la vista del mapa con algunos subgrupos de coordenadas obtenidos y sus centroides bajo el cálculo con distancias Euclidiana y Manhattan, donde se muestra que los subgrupos generados en ambas variantes tienden a una distribución espacial similar, aunque con diferente orden de los subgrupos. La minimización de estos grupos mediante un estimador, permitirá acercar el resultado a la orilla real. Por ejemplo, puede notarse que los centroides característicos de los subgrupos generados por el algoritmo se acercan y tienden a la orilla.

Figura 4.7. Subgrupos de coordenadas ruidosas obtenidos por *Kmeans*.



* Los otros subgrupos se han ocultado para efectos de presentación.

A partir de los grupos X_k generados por el algoritmo de *kmeans* partición espacial, se debe encontrar una operación Ψ que proporcione la coordenada más cercana a la orilla al

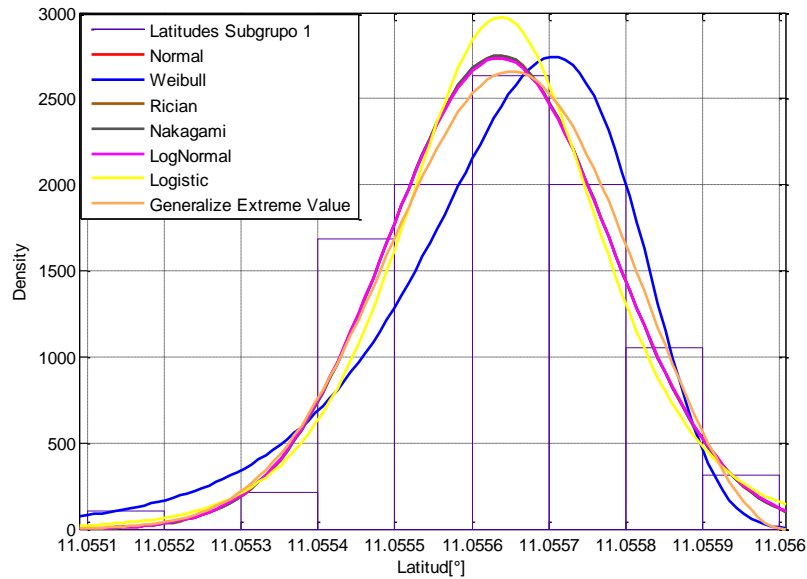
procesar los datos del subgrupo correspondiente, aplicando (4.2.4). La coordenada que se obtenga es en todo caso un parámetro o valor asociado con el proceso aleatorio de este subgrupo. Es decir, el problema de estimación que se debe resolver corresponde con la determinación del parámetro tridimensional $\langle \lambda_p \ \varphi_p \ h_p \rangle$ más cercano a la orilla obtenido del k -ésimo subgrupo.

Para simplificarlo, se debe reconocer que h_p es conocida debido a que las orillas están en el mismo plano geométrico de las coordenadas GPS de la embarcación y que las variables de longitud λ_p y latitud φ_p se pueden procesar de forma separada a partir de sus correspondientes muestras en cada subgrupo.

4.2.2.1 Aplicación de estimadores bayesianos a la reducción de incertidumbre en orillas generadas:

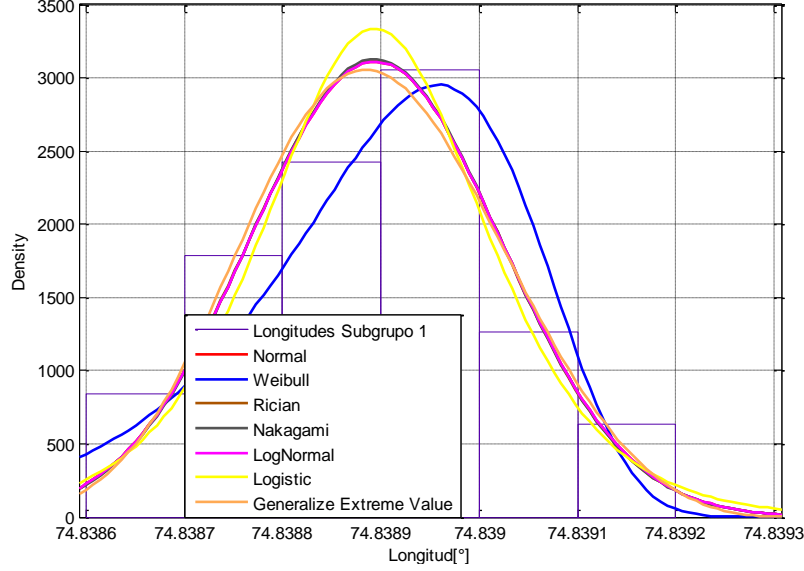
De acuerdo con la teoría bayesiana presentada en la sección 2.2.3, la función de distribución de salida $f_{Y|\Theta} (y |\Theta)$ hace parte de las condiciones para una estimación de riesgo reducido. Con miras a obtener la distribución de probabilidad de los subgrupos, al tomar alguno de estos y realizar un histograma por cada eje se obtienen resultados como el presentado en la Figura 4.8 y 4.9

Figura 4.8. Histograma de latitudes subgrupo 1 contra distintas distribuciones estadísticas.



De acuerdo con las Figura 4.8 y 4.9, los datos en cada eje se ajustan a una campana gaussiana típica de una distribución normal, forma que se repitió para todos los subgrupos evaluados. La verificación de la normalidad mediante el criterio χ^2 en una muestra aleatoria de veinte (20) subgrupos fue positiva.

Figura 4.9. Histograma de longitudes subgrupo 1 contra distintas distribuciones estadísticas.



En consecuencia, si el cálculo de los centroides por el método de distancias Euclidiana y *Manhattan* del algoritmo *kmeans* obtiene la media muestral y la mediana respectivamente (Mathworks, 2008), es de esperar que estos sean estimadores de máxima verosimilitud (MLE) y mínimo error absoluto (MAVE):

$$\hat{O}_{MLE}(k) = \bar{c}_{Euclidian}(k) : \text{Media muestral del } k\text{-ésimo subgrupo} \quad (4.2.8)$$

$$\hat{O}_{MAVE}(k) = \bar{c}_{Manhattan}(k) : \text{Mediana del } k\text{-ésimo subgrupo} \quad (4.2.9)$$

4.2.2.2 Estimador heurístico sesgado:

La revisión preliminar de resultados para los estimadores de media y mediana muestral que se mostrará en la sección 5.2, revela que si bien estos tienden a la media poblacional o el centro de la nube de muestras de cada subgrupo, la secuencia de coordenadas finales generalmente presentaba desviaciones sistemáticas, ubicándose geográficamente sobre tierra. Incluso, esta característica se presentó después de realizar la calibración de lecturas del radar con el procedimiento presentado en la sección 5.1.2.

Ante dicha situación, se hizo conveniente plantear la pregunta ¿Es el centro de la distribución o el centroide del subgrupo la coordenada más cercana a la orilla? ¿No será que en el proceso de detección, la tierra adentro genera mayor cantidad de detecciones del radar? ¿Por qué darle igual credibilidad a todas las muestras del grupo sabiendo que hay unas que fueron generadas con mayores rangos y por ende mayor incertidumbre?

Frente a estas consideraciones, se hace conveniente la evaluación de una función estimadora que busque la orilla así el resultado calculado por esta no sea la media del

proceso o el centro de la distribución de muestras. Es decir un estimador que pueda ser sesgado. Sin embargo, dada la complejidad intrínseca a la morfología de orillas, este sesgo más que un valor numérico general, representa una tendencia direccional variable en función de la ubicación del correspondiente subgrupo.

Se propone utilizar la información a priori, disponible de cada una de las muestras, por ejemplo la distancia de detección del radar $R(R_i)$ para la muestra en concreto) a la que una determinada coordenada fue generada. De tal forma que cada punto estimado de la orilla se obtuviera como:

$$\hat{O}_k = \Psi_k(X_k, R_i) \quad (4.2.10)$$

Donde la $\Psi_k(X_k, R_i)$ puede ser un promedio ponderado de las coordenadas del k -ésimo grupo en el que los pesos asignados a cada una dependan de la distancia a la que fue detectado. Luego, se tendrá que:

$$\hat{O}_k = \sum_{j=1}^{l_k} X_k(j) \cdot f_k(R_i, j) \quad (4.2.11)$$

Siendo f_k es una función de pesos que dependa del rango de detección de una determinada muestra y j un índice organizado de manera decreciente con las distancias de detección. Por ejemplo si un grupo está conformado por cuatro muestras, la media muestral implicaría asignar un peso de 0.25 a cada una en el promediado; Bajo el nuevo enfoque, si estas fueron generadas por una detección a 250 m, 300 m, 330 m y 450 m se les podría asignar como pesos 0.5, 0.25, 0.15 y 0.1 respectivamente llevando a darle más influencia a la muestra de menor distancia de detección.

Para el caso presente, la intención de asignar estos pesos resulta una operación complicada, debido a que hacemos referencia a subgrupos de $K = 50, 120, 155$, etc., en general cantidades grandes y arbitrarias de elementos lo que dificulta asignar esa cantidad de pesos de tal forma que se mantenga la unidad. La función f_k que se utilice deberá ser de carácter decreciente, de tal forma que la suma de sus términos sea la unidad. La serie geométrica de la forma expuesta a continuación, cumple esta característica:

$$f_l(R_i) = \frac{\beta}{(\beta + 1)^l} \quad \beta = 0.1, 0.2, 1, 2, 3, \dots \quad (4.2.12)$$

Donde la dependencia con respecto a la distancia de detección de la muestra R_i , se genera mediante la asignación decreciente de los pesos f_l a cada coordenada en el subgrupo correspondiente. Es decir la coordenada generada por la distancia más corta

tendrá el peso más alto dado por $\beta/(\beta+1)$ y la coordenada del subgrupo con distancia de detección más lejana se la asignará $\beta/(\beta+1)^l$. El estimador propuesto se definirá consecuentemente como:

$$\hat{O}_{beta}(k) = \sum_{j=1}^{l_k} X_k(j) \cdot \frac{\beta}{(\beta+1)^j} \quad (4.2.13)$$

4.3 MÉTRICAS DE DESEMPEÑO PARA EL SISTEMA

La toma de decisiones y generación de conclusiones con respecto a las técnicas de estimación propuestas y sus variantes dependerá de que se pueda establecer unas medidas de calidad o desempeño generales para la secuencia final de coordenadas obtenidas.

Se proponen como métricas de desempeño de la estimación realizada en 4.2.1 y 4.2.2 a cantidades relacionadas con la resolución y el error. La resolución se constituye como un indicador de calidad típico de sistemas con adquisición remota como tomas satelitales, aéreas y radares.

Sobre el error, se define de manera bidimensional, en el eje Norte–Sur y eje Este–Oeste medido con respecto a una secuencia de coordenadas de referencia. Para el error en cada eje se calculan estadísticos debido al carácter aleatorio de estos en cada coordenada de muestra, siendo de común utilización el máximo (*MAX*), mínimo (*MIN*), la raíz cuadrada del valor medio (*RMS*) y finalmente el valor por debajo del cual se encuentran el 95% de las desviaciones, que corresponde con la desviación estándar. Al computar las desviaciones estándar de ambos ejes, se obtiene la métrica estadística bidimensional Error Circular Probable (ECP), que representa el radio del círculo dentro del cual se encuentran el 95% de las muestras arrojadas para una estimación de coordenadas, medida de calidad frecuentemente utilizada en receptores GPS.

Los estadísticos del error que se han descrito, tienen aceptación como indicadores de calidad de sistemas de medición de líneas de orilla y posicionamiento preciso por GPS, como puede verificarse en los trabajos relacionados en (Marfai y otros, 2008) y (Harley, 2010).

4.3.1 Cálculo de la Resolución

Corresponde con la distancia media o separación entre las coordenadas de la orilla generadas como solución. Note que esta cantidad se relaciona con la capacidad del sistema para describir la morfología de las orillas, en tanto un valor de resolución alto implicará que algunos detalles geográficos no puedan describirse adecuadamente. La *Resolución* se calculará como la distancia media intra-coordenadas para la secuencia de

coordenadas de la solución, buscándose la coordenada más cercana a cada una de estas. El parámetro *Resolución* que se denotará como Δ , se calcularía como:

$$\Delta = \frac{1}{(K-1)} \sum_{k=1}^K d_E(\hat{O}_k, \text{Closest}(\hat{O}_k, \hat{O})) \quad (4.2.14)$$

Donde se ha requerido en la definición a los operadores definidos en la sección **4.2.1**, debido a que la solución se presenta como un conjunto de coordenadas, más no una secuencia ordenada, de ahí que el cálculo del promedio implique tomar cada coordenada y determinar su distancia hasta la coordenada más cercana del conjunto, acumulando este valor y realizar el procedimiento para cada una, para finalmente dividir entre el total de distancias sumadas. El proceso a nivel algorítmico debe incluir previsiones para evitar la repetición de distancias y para los extremos de la secuencia.

4.3.2 Cálculo de estadísticos del error

El error debería medirse con respecto a la línea de orilla real. Sin embargo, las medidas de error a utilizar se establecerán a partir de las distancias geográficas entre cada una de las K coordenadas estimadas con respecto a las K coordenadas más cercanas de un conjunto O_{ref} que describe la ubicación casi exacta o de referencia para la orilla.

Al establecer la referencia, la magnitud del error notada como $e(k)$ para cada subgrupo se calcularía como:

$$e(k) = \text{Convert}\{d_E(\hat{O}(k), \text{Closest}(\hat{O}(k), O_{ref}))\} \quad (4.2.15)$$

Donde se han utilizado en la definición operadores definidos en la sección **4.2.1** y se ha incluido la operación $\text{Convert}(\cdot)$ para representar que el valor obtenido se debe transformar a metros si se utilizan coordenadas elipsoidales (WGS84) o convertirlas al sistema cartesiano rectangular. Para el ángulo del error $\theta_e(k)$, utilizando la relación del producto punto con respecto a un vector al norte en la proyección cartesiana (0,1,0) se tiene:

$$\theta_e(k) = \cos^{-1} \left(\frac{\hat{O}(k) - \text{Closest}(\hat{O}(k), O_{ref}) \bullet (0,1,0)}{d_E(\hat{O}(k), \text{Closest}(\hat{O}(k), O_{ref}))} \right) \quad (4.2.16)$$

Luego, se obtendrá el error para los ejes Norte - Sur y Este - Oeste respectivamente:

$$e_N(k) = e(k) \cdot \cos(\theta_e(k)) \quad (4.2.17)$$

$$e_O(k) = e(k) \cdot \text{sen}(\theta_e(k)) \quad (4.2.18)$$

Luego, los estadísticos del error se determinarían aplicados en cada eje como:

- El máximo error registrado en las K coordenadas finales de estimación:

$$e_{MAX} = \arg_{e(k)} \max[e(k)] \quad (4.2.19)$$

- El mínimo error registrado en las K coordenadas finales de estimación:

$$e_{MIN} = \arg_{e(k)} \min[e(k)] \quad (4.2.20)$$

- El error absoluto medio o raíz del cuadrático medio registrado para las K coordenadas finales de estimación:

$$e_{RMSE} = \frac{1}{K} \sum_{k=1}^K \sqrt{(e(k))^2} \quad (4.2.21)$$

- Finalmente, el Error Circular Probable se obtiene incluyendo la desviación estándar σ en cada eje, de tal forma que:

$$\sigma_N^2 = \frac{1}{K} \sum_{k=1}^K (e_N(k) - \bar{e}_N)^2 \quad (4.2.22)$$

$$\sigma_O^2 = \frac{1}{K} \sum_{k=1}^K (e_O(k) - \bar{e}_O)^2 \quad (4.2.23)$$

$$e_{ECP} = \sqrt{\sigma_N^2 + \sigma_O^2} \quad (4.2.24)$$

5 RESULTADOS OBTENIDOS Y ANÁLISIS

A continuación se exponen los resultados y análisis obtenido a través del proceso investigativo, buscando mostrar la evolución del producto final, en función de las variantes evaluadas y las métricas de desempeño. Los resultados se han desarrollado a partir de datos tomados en dos secciones de campo, realizadas en los días 12 y 19 de julio de 2010. Se escogió para las pruebas un sector del Río Magdalena aledaño al barrio las Flores de Barranquilla y la desembocadura de este río Bocas de Ceniza. La escogencia de dicho lugar se realizó por estar ubicada en cercanías de donde se ha desarrollado esta investigación y que el sector cumple con la característica de tener orillas muy estables debido a morfología meándrica y algunas obras de canalización como espolones y diques direccionales en la zona.

En este sentido, se utilizarán para comparación los resultados que se obtengan del inicio del tamar occidental de Bocas de Ceniza, que podrán ser validados mediante el sistema de información geográfico Google Earth y archivos de levantamiento cartográfico .shp generados por el Instituto de Estudios Hidráulicos y Ambientales (IDEHA) de la Universidad del Norte. Los resultados numéricos y la mayoría de gráficas se realizaron por programación en Matlab, donde también se importó el correspondiente archivo .shp del IDEHA y coordenadas de la orilla en el sector extraídas de Google Earth.

El desempeño se determinará en función de las métricas definidas en la sección 4.3, donde se estableció que serían la resolución y algunos estadísticos del error. Para la obtención de la resolución se calcula la distancia promedio interna del vector de coordenadas de la solución de acuerdo con lo expuesto en la sección 4.3.1. Para el cómputo de los estadísticos del error, sobre las K coordenadas de solución y tomando vector de referencia para orillas “exactas” O_{ref} , de una secuencia de coordenadas rigurosamente importadas desde Google Earth, se aplican los procedimientos y fórmulas expuestas en la sección 4.3.2, donde se definió:

- RMS: Raíz del error cuadrático medio, coincidente con el error absoluto promedio.
- MAX: Máximo valor de error absoluto registrado en un eje.
- MIN: Mínimo valor de error absoluto en un eje.
- STD: Desviación estándar o valor por debajo del cual se encuentran el 95% de errores.
- ECP: Relación pitagórica de las desviaciones estándar de los dos ejes, obteniéndose una superficie radial dentro de la que se encontrará el 95% de errores en coordenadas.

5.1 RESULTADOS PRELIMINARES DEL SISTEMA SIN ESTIMACIÓN

Antes de obtener los resultados definitivos del sistema, se presentan las coordenadas generados al aplicar el procesamiento de las entradas de radar, GPS y sensor de rumbo de los algoritmos de estimación, con la finalidad de identificar incoherencias y errores sistemáticos que se puedan estar presentando.

5.1.1 Calibración de Mediciones del Radar

Ante los resultados preliminares de coordenadas generadas, se determinó necesario un proceso de calibración de las mediciones de distancia de radar frente a algún error sistemático. Se hizo conveniente en consecuencia, encontrar el posible valor de desviación del radar para corregir todas las distancias entregadas por este. Para ello, se seleccionó una boya por ser un elemento geográfico lo suficientemente separado del resto como para poder identificar en Google Earth la distancia y azimuth desde las cuáles pudo haber sido capturado por el radar. Además, que su coordenada geográfica se encuentra disponible en la información cartográfica provista por el IDEHA.

Luego, las distancias de Google Earth se fijaron como referencia “exacta” frente a las distancias registradas por la segmentación de orillas en la imagen de radar establecidas en el ángulo dado por la diferencia entre el azimuth de google Earth y el azimuth del sensor de rumbo correspondiente a esa toma. La Figura 5.1 ilustra cómo se tomaron las distancias y azimuths de referencia con la herramienta “Regla” de Google Earth. El contraste de distancias de detección entre el radar y google Earth se presenta en la Tabla 5.1.

Figura 5.1. Distancias y ángulos de detección de la Boya de referencia para calibración, desde cada posición GPS de la embarcación.

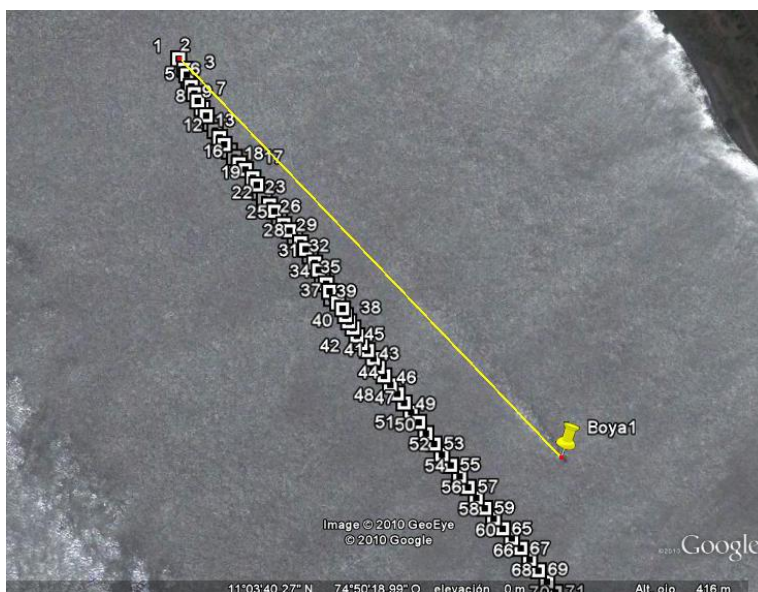


Tabla 5.1. Datos que relacionan distancias desde Google Earth y distancias entregadas por Radar

Posición GPS	Longitud [°]	Latitud [°]	Distancia Google Earth	Azimut Google Earth [°]	Distancia radar [m]	Azimut Radar [°]	Error [m]
1	-74,83981388	11,062266	345,752011	136,023153	360,2253521	136,023153	14,4733407
6	-74,83972353	11,0620703	323,413327	134,618122	336,7323944	137,618122	13,3190676
7	-74,83970353	11,0620293	318,676112	134,313541	332,8169014	134,313541	14,1407898
9	-74,83965503	11,0619458	308,449432	133,77432	317,1549296	130,77432	8,70549722
11	-74,83960383	11,0618626	298,055136	133,240712	309,3239437	133,240712	11,2688079
12	-74,83957778	11,0618233	293,006808	133,002915	301,4929578	133,002915	8,48614997
13	-74,8395516	11,0617844	287,984066	132,765328	305,4084507	131,765328	17,4243843
15	-74,83949179	11,061712	277,747105	132,466685	293,6619719	129,466685	15,9148665
20	-74,83934555	11,0615162	251,391289	131,28429	262,3380282	131,28429	10,9467388
23	-74,83926182	11,0614096	236,757645	130,601983	250,5915493	130,601983	13,833904
26	-74,83917523	11,0613004	221,733553	129,823234	234,9295775	126,823234	13,1960248
29	-74,83908178	11,0611962	206,522105	129,181276	219,2676056	129,181276	12,7455011
60	-74,83804923	11,0597409	56,2522854	57,1624971	70,47887324	57,1624971	14,2265879
67	-74,83769137	11,0593287	76,5352688	6,12196601	86,14084508	6,12196601	9,60557626

De las muestras realizadas se obtuvo que la desviación promedio en los rangos mantiene un valor sistemático de 12.73 m. En consecuencia, para efectos de este trabajo cada una de las distancias de detección medidas con el radar fueron ajustadas restándoles 12.73 m.

5.1.2 Georreferenciación Inicial con Radar

De la sesión de campo del 19 de julio se recopilieron 254 tomas de radar (imágenes) adjuntas a 254 frames de datos de GPS y sensor de azimut. Estos datos se pasaron por el algoritmo de georreferenciación, generando una coordenada por cada punto de la orilla detectado con radar. La salida del sistema genera nubes de abundantes puntos dispersos con respecto a la línea de orilla buscada, situación que se ilustra en la Figura 5.2., donde se ha utilizado Google Earth de apoyo para efectos de visualización.

De estos resultados parciales, a pesar de la dispersión y redundancia propias de esta etapa del proceso, se verifica la consistencia del sistema en cuanto ha generado información geográfica muy relacionada con las orillas buscadas, siendo las imprecisiones causadas por los equipos electrónicos de entrada.

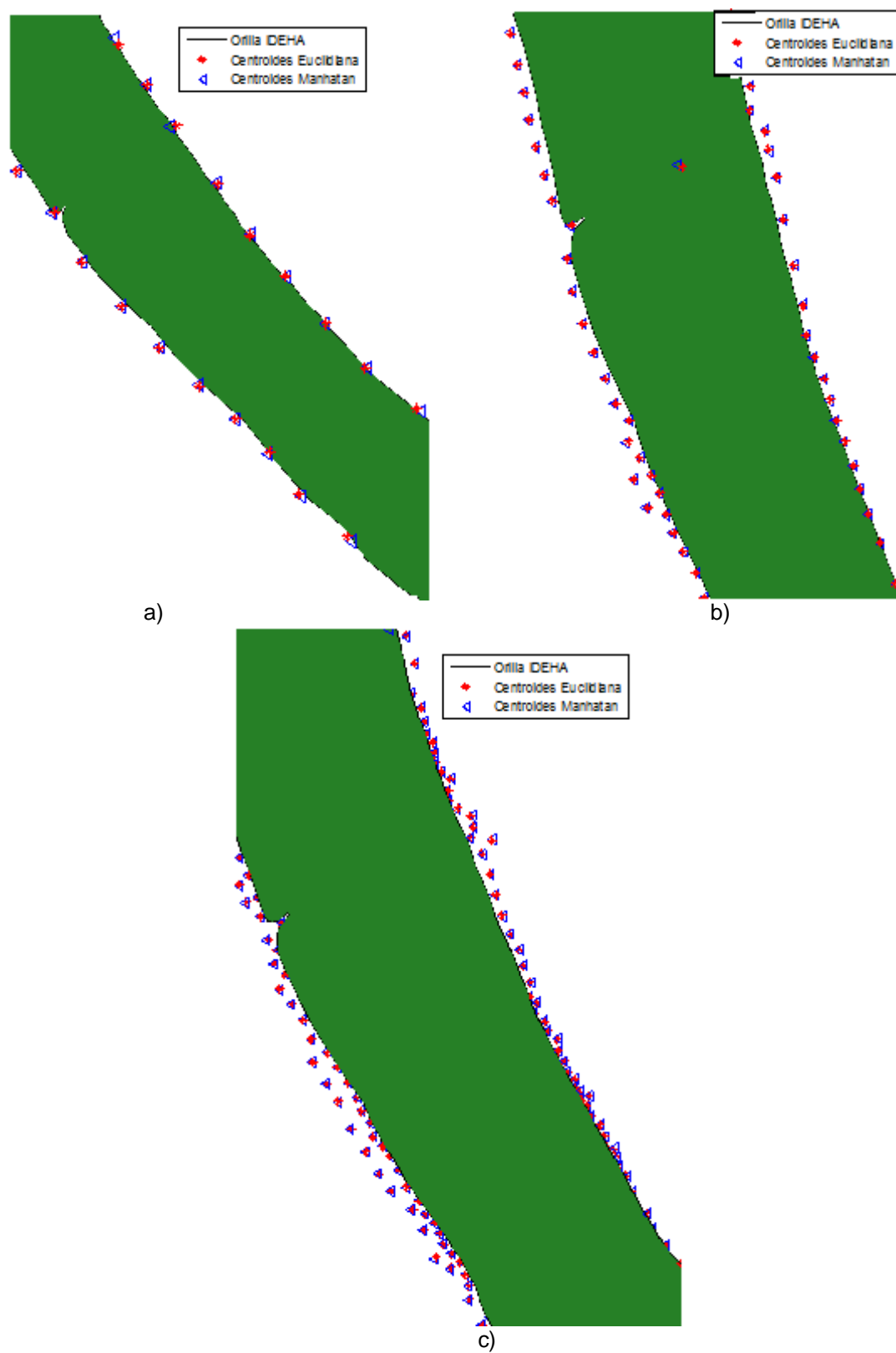
Figura 5.2. Coordenadas ruidosas generadas por el sistema hasta el algoritmo de georreferenciación.



5.2 OPTIMIZACIÓN DE LA CANTIDAD DE SUBGRUPOS GENERADOS

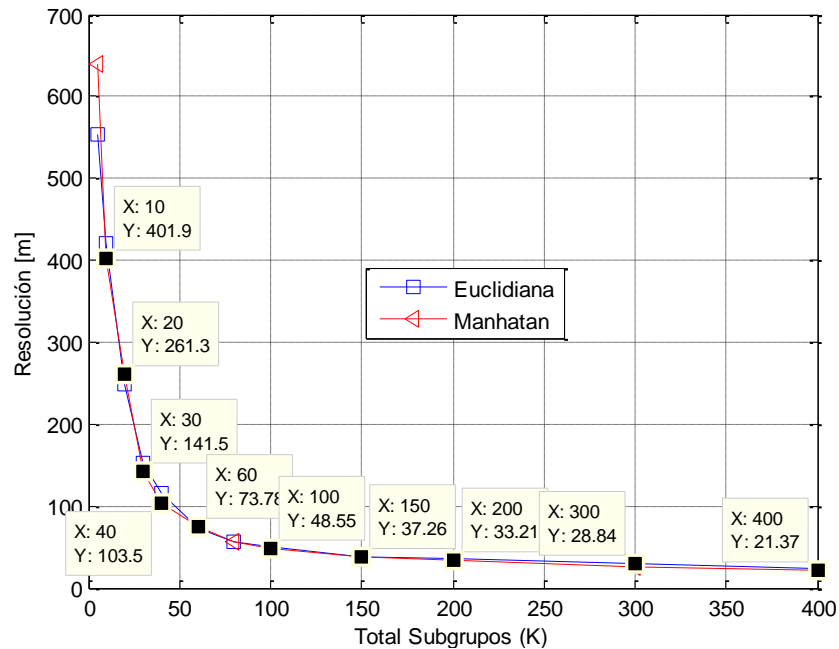
Dada la escogencia del algoritmo *Kmeans* para la segmentación de los puntos significativos de la orilla se hace necesario contrastar las variantes de distancia Euclidiana y *Manhattan* implementadas. En este caso se hizo la comparación variando la cantidad K de subgrupos a generar y revisando el desempeño obtenido, además de la inspección visual de los resultados en el mapa. A nivel gráfico se verificó un adecuado comportamiento de los subgrupos en cuanto tendían a concentrarse en cercanías de la orilla, como se presenta en la Figura 5.3. Para bajos valores de K los puntos no logran definir algunos detalles de la geografía. Para valores de K altos, la complicación se origina por la aparición de algunos centroides con alta desviación, los cuales incrementan el error del sistema.

Figura 5.3. Estimación de coordenadas de orilla mediante centroides del *kmeans*



De la Figura 5.3, a) corresponde a los centroides generados para $K=20$, b) los centroides para $K=60$ y c) los centroides para $K=150$. Sin embargo, los patrones espaciales no proporcionan información cuantitativa para descartar a alguna de las variantes del algoritmo *kmeans*, de ahí que se planteó el experimento de evaluarlo en la generación de $K = 5, 10, 20, 30, 40, 60, 80, 100, 150, 200, 300$ y 400 subgrupos. En la Figura 5.4 se presenta la relación entre el número de subgrupos y la resolución del sistema para las optimizaciones de distancia Euclidiana y Manhattan. Al revisar dicha figura se verifica que la resolución presenta un comportamiento similar para la optimización Euclidiana y Manhattan, confirmando lo presentado en la Figura 5.3, dado que ambas tienden a generar subgrupos sin diferencias apreciables.

Figura 5.4. Relación de resolución del sistema obtenida para el *Kmeans* en las variantes implementadas.

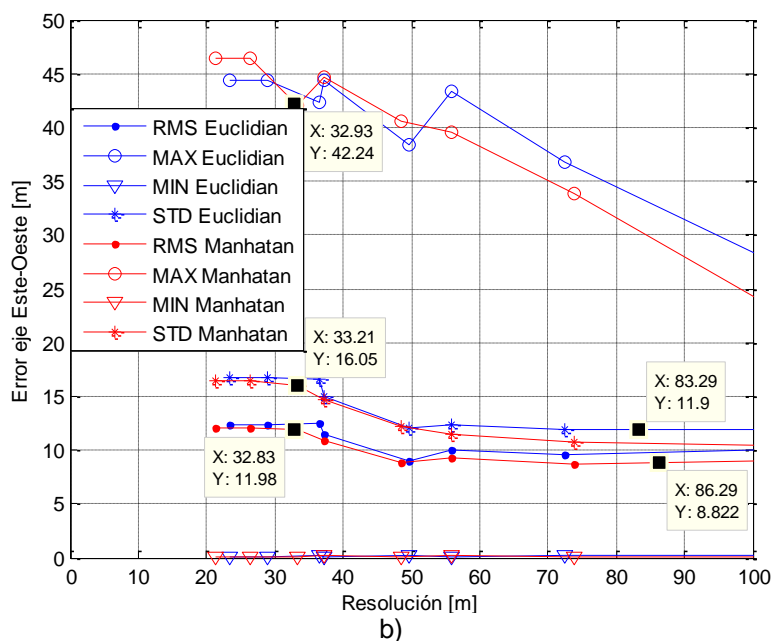
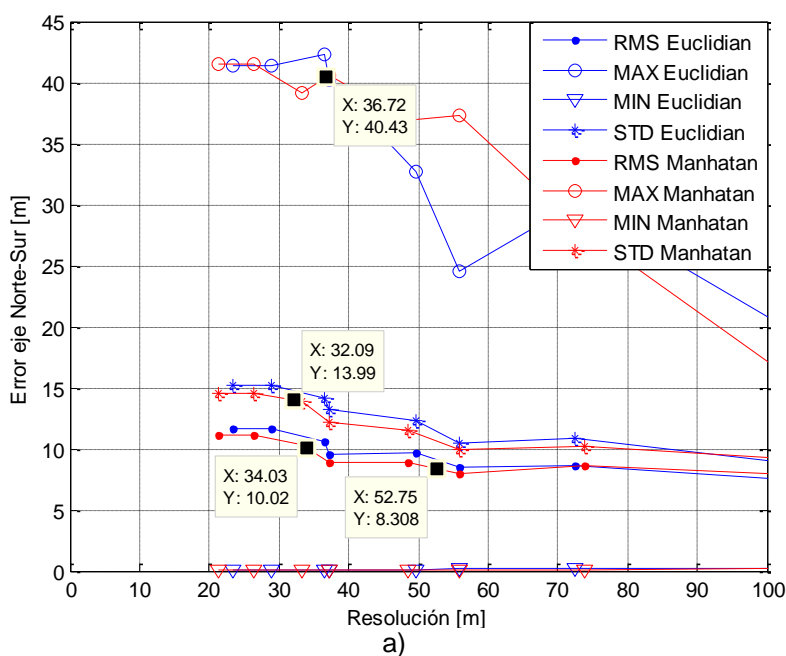


Adicionalmente, se encontró un efecto de saturación en la caída de la resolución, pues se llega a un valor de K a partir del cual el incremento del número de subgrupos no ocasiona una disminución importante en la resolución. La situación ocurre porque la aparición de nuevos subgrupos se presenta por fuera de la línea de orilla (ver Figura 5.3c)), que se alejan de las coordenadas adyacentes incrementando el error del sistema. El punto de saturación de la resolución se presentó a partir de cien (100) subgrupos totales. Si bien el valor mínimo de resolución ($\sim 20\text{m.}$) se presentó para $k=400$ es conveniente revisar que sucede con el error.

Al utilizar la relación encontrada en la Figura 5.4 entre la cantidad de subgrupos y la resolución, ésta se contrasta frente a las estadísticas del error en cada eje, resultados mostrados en la Figura 5.5, donde se ha hecho énfasis en el rango de resolución por

debajo de 100 m. dado que valores por encima representan una calidad inaceptable para la mayoría de aplicaciones en contraste con otras técnicas de georreferenciación.

Figura 5.5. a) Relaciones de error en el eje Norte-Sur obtenido para el sistema frente a la resolución. **b)** Relaciones de error en el eje Este-Oeste obtenido para el sistema frente a la resolución.



A partir de los resultados mostrados en las Figura 5.4 y 5.5, se pueden realizar algunas conclusiones.

- La primera es que el sistema de georreferenciación propuesto en esta tesis presenta una resolución mínima de 20 m., límite correspondiente a $K=400$ bajo las condiciones en que se realizó este trabajo.
- La variación del error en su valor RMS y STD no es muy pronunciada con respecto a la resolución, a diferencia de los errores máximos que descienden bruscamente a medida que el valor se aumenta. Los valores RMS y STD del error están siempre por debajo de 20 m.

Teniendo en cuenta la tendencia de desempeño del sistema al contrastar resolución contra medidas del error, se considerará una resolución alrededor de los 50 m., debido a que a partir de este punto no hay disminuciones sustanciales en los valores RMS y STD. Además, considerando que las mediciones del IDEHA manejan un valor de resolución cercano a los 30 m. con espaciamientos hasta de 70 m. en las partes rectas de la geografía y hasta de 10 m. en curvas y elementos que demandan mejor resolución, según las medidas realizadas en un archivo .shp suministrados por este grupo de investigación.

De acuerdo con esta resolución de referencia y la curva del sistema frente a la cantidad de subgrupos (Véase Figura 5.4), debe ajustarse a $K=100$ para las dimensiones geográficas del sector escaneado. La selección está justificada además porque a partir de este valor empiezan a aparecer coordenadas erráticas (*outliers*) producto de centroides que no tienden a la orilla, como se puede constatar en la Figura 5.3c) para $K=150$, donde los *outliers* se presenta con una frecuencia y tendencia significativa.

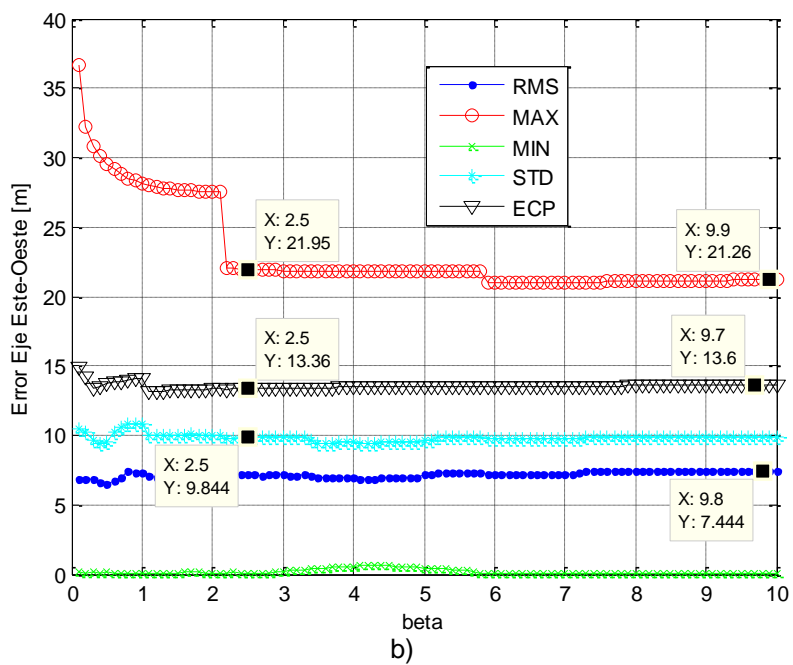
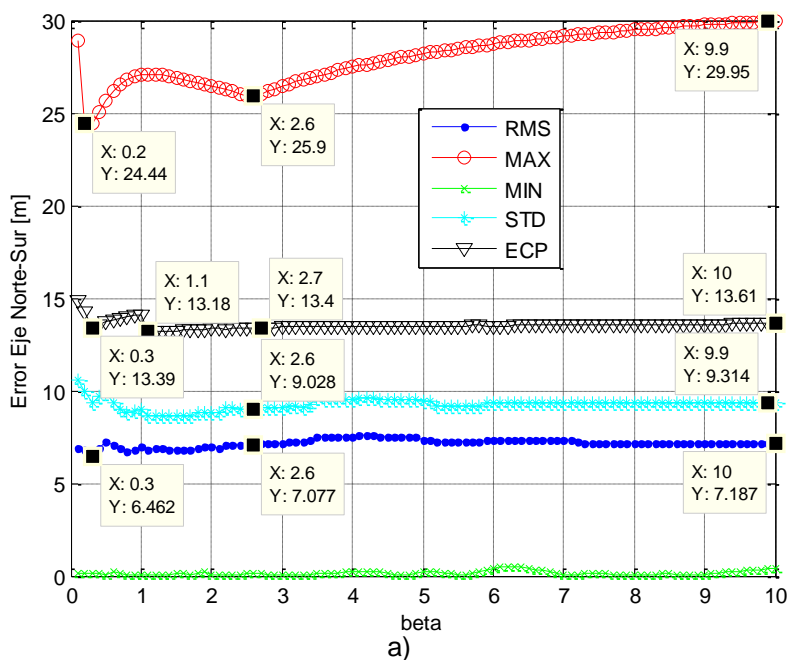
5.3 SELECCIÓN DEL ESTIMADOR PARA REDUCCIÓN DE INCERTIDUMBRE

De los anteriores resultados se escogió un valor de $K=100$ para mantener el error reducido. A continuación se escogerá entre las técnicas de reducción de incertidumbre a los centroides de distancia Euclidiana que genera la media, los centroides de distancia Manhattan que genera la mediana y el estimador sesgado propuesto en la sección 4.2.2. Cabe retomar que dichos estimadores, de acuerdo con lo presentado definen tres funciones distintas Ψ_{MLE} (media), Ψ_{MAVE} (mediana) y Ψ_{BETA} (Estimador Sesgado propuesto) para la implementación de la expresión (4.2.4).

Antes de la comparación, dada la dependencia de la función Ψ_{BETA} con el parámetro β se calcularon curvas de desempeño en cada métrica, resultado que se presenta en la Figura 5.6a) y 5.6b). Se encuentra principalmente que los estadísticos promediadores se mantienen constantes ante el parámetro β a diferencia del error máximo que varía de distinta manera por intervalos y por eje. Este comportamiento probablemente se presente

por el efecto del estimador sobre algunos *outliers* provocando que estos se desplacen en una determinada dirección.

Figura 5.6. a) Comportamiento del error en eje Norte-Sur en función de beta. **b)** Comportamiento del error en eje Este-Oeste en función de beta.



Para la selección de β al considerar el carácter casi constante de los otros estadísticos, se haría basado en el error máximo, según el cual, el valor de menor riesgo se obtendrá para $\beta = 2,5$. Nótese que el error máximo en el eje Norte-Sur presenta un mínimo local en ese valor (25,9 m), a partir del cual aumenta hasta 30 m. Por parte del eje Este-Oeste, el error se mantiene constante en 21,95 m., después de haber descendido bruscamente en $\beta = 2,1$. En consecuencia, para efectos de resultados definitivos, el estimador sesgado propuesto se utilizará con $\beta = 2,5$.

En la Tabla 5.2 se presenta el desempeño del sistema de georreferenciación en las variantes de estimación implementadas frente a los estadísticos del error y la resolución. De estos resultados se encuentra que el estimador propuesto mejora en cada una de las métricas a la estimación de los centroides por media y mediana, lo que lo hace idóneo para presentación final del sistema ante otras técnicas de georreferenciación. En especial se encuentra una reducción importante en los errores máximos registrados, lo que a su vez influye sobre los otros estadísticos del error.

Tabla 5.2. Desempeño de variantes del sistema

Estimador	Error	MAX [m]	MIN [m]	RMS [m]	STD [m]	ECP [m]	Resolución
Euclidiano (Media)	Norte-Sur	32,67	0,04	9,67	12,25	17,15	57,22
	Este-Oeste	38,30	0,29	9,04	12,00		
Manhattan (Mediana)	Norte-Sur	36,86	0,04	8,85	11,45	16,68	56,88
	Este-Oeste	40,56	0,07	8,84	12,13		
Propuesto ($\beta = 2,5$)	Norte-Sur	21,95	0,09	7,06	9,02	13,35	50,04
	Este-Oeste	25,95	0,02	7,20	9,84		

Por otro lado, un resultado no esperado fue una mejoría (disminución) significativa en la resolución, permitiendo afirmar que los *outliers* en la solución final no sólo empeoran el error sino también a la métrica de resolución debido a la aparición de mayores distancias intra coordenadas hacia las muestras erráticas.

5.4 DESEMPEÑO FRENTE A OTRAS TÉCNICAS DE GEORREFERENCIACIÓN

A través del proceso de las secciones anteriores, se ha conseguido un sistema de georreferenciación con el desempeño presentado para el estimador sesgado propuesto. A continuación se establecerá la comparación con respecto a la técnica de georreferenciación de levantamientos directos con GPS de alta precisión, utilizada a nivel estándar por el IDEHA en sus trabajos de batimetría y topografía fluvial en investigaciones y proyectos de consultoría.

En la Figura 5.7 se presentan las secuencias de orilla finales obtenidas por el sistema de georreferenciación propuesto, superpuestas a las secuencias de coordenadas del mismo sector obtenidas por Google Earth y el IDEHA. La ampliación correspondiente en algunos sectores es presentada en la Figura 5.8, donde se enfatizaron los tajamares de del canal de acceso de Bocas de Ceniza y se ha realizado un giro de -90° sobre las imágenes para efectos de presentación.

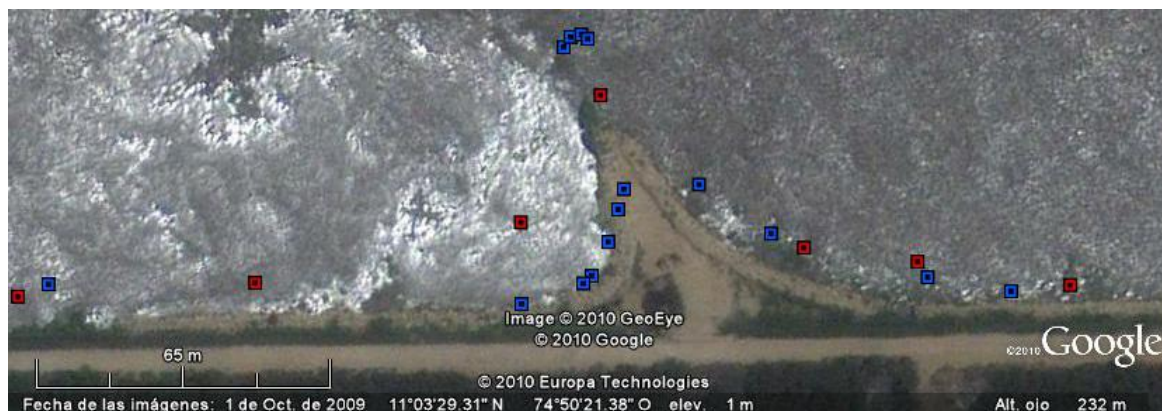
Figura 5.7. Coordenadas de orillas estimadas por Sistema y por IDEHA sobre Google Earth.



*Cuadros azules: coordenadas solución IDEHA. Cuadros rojos: coordenadas estimadas por sistema propuesto.

De las Figuras citadas, se puede afirmar que la propuesta de georreferenciación propuesta puede describir la geografía fluvial de manera exitosa, sin pérdida importante de detalles que pudieran afectar la navegación, tales como espolones. La revisión cuantitativa conlleva a la revisión de ambos sistemas en función de las métricas de desempeño planteadas en este trabajo, lo cual se presenta en la Tabla 5.3.

Figura 5.8. a) Ampliación de orillas finales obtenidas (sector de tajamar occidental. b) Ampliación de orillas finales obtenidas (sector de tajamar oriental).



a)



b)

*Cuadros azules: coordenadas solución IDEHA. Cuadros rojos: coordenadas estimadas por sistema propuesto.

Tabla 5.3. Desempeño del sistema de georreferenciación

Sistema	Error	MAX [m]	MIN [m]	RMS [m]	STD [m]	ECP [m]	Resolución
IDEHA*	Norte-Sur	24,83	0,04	5,35	7,31	10,08	31,42
	Este-Oeste	21,52	0,18	4,93	6,95		
Propuesto**	Norte-Sur	21,95	0,09	7,06	9,02	13,35	50,04
	Este-Oeste	25,95	0,02	7,20	9,84		

* Levantamientos directos con GPS alta precisión.

** Estimación remota mediante procesamiento y transformación de radar de bajo costo.

De los resultados consignados, se encuentra que el desempeño del sistema por levantamientos directos con GPS es superior, aunque como se planteó en la problemática de esta investigación presente un mayor costo de elaboración. Aún así, el desempeño en las métricas del error confirma la posibilidad de implementar el sistema para levantamiento de información geográfica de orillas en cartas náuticas, teniendo en cuenta que el ECP típico en los GPS's de navegación se encuentra alrededor de los 15 m.

6 CONCLUSIONES

Con la presentación de esta memoria se ha expuesto una síntesis del trabajo desarrollado en el marco del cumplimiento de los objetivos de investigación para la tesis de maestría del autor, con lo que se han obtenido importantes productos o resultados que se deben anotar:

- Una nueva técnica de aplicación de radar en sensado remoto, en este caso específica para la recolección de información de coordenadas de orilla en canales fluviales. La propuesta al ser demostrada en su viabilidad técnica podrá ser utilizada para disminuir los costos de actualización para la cartografía fluvial en sectores trezados.
- Un modelo matemático para cálculo de coordenadas proyectándolas con azimut y radio, desarrollado a partir de las curvas paramétricas tridimensionales que describen el elipsoide terrestre en toda su superficie, lo que permitió una solución generalizada a cualquier coordenada geográfica y extensible a otros contextos como la corrección de errores sistemáticos en sistemas de información geográfica.
- Se realizó la implementación real del sistema, lo cual se demostró con los datos reales recopilados y procesados mediante el conjunto de desarrollos de ingeniería expuestos en la sección 3, en los que se creó un sistema de adquisición para tomas de radar marino sincronizada a sentencias NMEA en dispositivos como GPS, sensor de azimut, desarrollo que puede ser registrado en su modalidad de software o desarrollo industrial.
- Se implementó una estrategia de estimación que aprovechó la redundancia de datos ruidosos del proceso y generó un conjunto reducido y de mayor exactitud de las coordenadas de la orilla, obteniendo una resolución de 50 m., un Error Circular Probable (ECP) de 13.35 m.. La estrategia de estimación basada en dos etapas: la primera un algoritmo de agrupación por *kmeans* y la segunda con la aplicación de estimadores bayesianos para reducir cada subgrupo a una coordenada.
- Se propuso un estimador inédito en la literatura que introduce sesgo a favor de las coordenadas de subgrupo con menor distancia de detección del radar, mejorando a los estimadores MLE (media) y MAVE (Mediana) para efectos de la aplicación, debido a la introducción de mayor conocimiento del proceso.
- Se presentaron y explicaron métricas de desempeño en la sección 4.3 que sirvieron para comparar las variantes del estimador, la referencia de coordenadas provistas por el Sistema de Información Geográfico por procesamiento de imágenes remotas de Google Earth y las coordenadas suministradas por el IDEHA para el sector evaluado.

- Se validó el sistema en un contexto real, realizando la georreferenciación de un segmento de orillas en Bocas de Ceniza, Barranquilla. Los resultados de dicha intervención fueron comparados con dos sistemas alternos: Google Earth y levantamientos directos con GPS por parte del IDEHA, permiten afirmar que la exactitud obtenida para el sistema lo hace apto para navegación.

RECOMENDACIONES

En un proceso investigativo como el que se ha expuesto siempre quedan en el camino alternativas y mejoras que no pueden abordarse porque realmente se excedería el alcance del proyecto en un determinado campo y que extendería el tiempo de desarrollo a tal punto de que se corra el riesgo de pérdida de vigencia en el problema a resolver. Al finalizar esta tesis, se propone asumir el trabajo realizado como una primera etapa, que debería continuarse mediante las siguientes líneas de acción.

Mejorar el desempeño de error y resolución del sistema, enfatizándose en reducir la incertidumbre en las mediciones de radar no solo con la utilización de un dispositivo más preciso sino revisando el esquema de procesamiento de las tomas de radar, buscando disminuir posibles fuentes de errores en la captura de la imagen. Por otro lado, se debería caracterizar experimentalmente el radar empleado a nivel de sus características de incertidumbre en la medida entregada.

Realizar trabajos exploratorios para lograr la medición de distancias a orilla con ecosonda, que podría aportar la precisión bajo agua de estos dispositivos como referencia para calibración del error sistemático y varianza en las mediciones de radar. Esto se conseguiría con la ecosonda reorientada hacia un modo horizontal por debajo del agua, con la finalidad de medir distancias a orillas desde la embarcación en un ángulo de 90° con respecto a la proa. Esta iniciativa tiene el potencial de mejorar considerablemente la exactitud del sistema, en cuanto las mediciones de ecosonda, en modo normal tienen un error del orden de centímetros.

Explorar las variantes y algoritmos disponibles en el estado actual de desarrollo para la partición de los subgrupos de coordenadas del proceso, incluyendo técnicas de inteligencia artificial como el *kmeans* difuso y algoritmos de segmentación bayesiana para grupos gaussianos entremezclados (*Gaussian Mixture Models*). En esta parte se podrían estudiar mecanismos y estrategias para incorporar en la estimación final la información a priori de las coordenadas redundantes de tal manera que se disminuya los errores obtenidos.

Evaluar este sistema de sensado remoto en otras aplicaciones como el levantamiento topográfico de playas, abriéndole un mayor potencial de uso.

REFERENCIAS Y BIBLIOGRAFÍA

ALVARADO, M. Sistema de Asistencia Satelital a la Navegación en el Río Magdalena entre Puerto Salgar (K925) y Barranquilla. Bogotá: Conferencia Feria Internacional del Libro, 2009.

ARMSTRONG, N. & Hibbert, D. An introduction to Bayesian methods for analyzing chemistry data Part 1: An introduction to Bayesian theory and methods. Science Direct: Chemometrics and Intelligent Laboratory Systems, 2009.

CASTRO, V. & JANG, J. Fast and Robust General Purpose Clustering Algorithms. Kluwer Academic Publishers: Data Mining and Knowledge Discovery, 2004.

CHEN, X. Study on FOG-SINS/GPS integrated attitude determination system using adaptative Kalman filter. Springer-Verlag Berlin Heidelberg, 2006. 417-427p.

CIAMPA, Jhon. Method and apparatus for mapping and measuring land, patent number 5247356. United States Patent, 1993.

MINISTERIO DE TRANSPORTE (Colombia). “Manual de Ríos Navegables”. Bogotá, 2000.

CRACKNELL A. & HAYES L. Introduction to Remote Sensing. 2 ed. 2007. 21-71p.

DANKERT, H. Retrieval of Surface-Current Fields and Bathymetries using Radar-Image Sequences. IEEE. 2003.

DANKERT H. & ROSENTHAL, W.. Ocean surface determination from X-band radar-image sequences. Journal of Geophysical Research, Vol. 109. 2004.

DATTA, A. & DOUGHERTY, E. Introduction to Genomic Signal Processing with Control. CRC Press. 2007. 167–180p.

DENINSON D.G.T. & LAWSON A.B., Spatial Cluster Modelling, Chapman and Hall/CRC. 2002. Chapter 1.

DIRECCIÓN NACIONAL DE PLANEACIÓN DNP (Colombia). Documento “Visión Colombia II Centenario”, capítulo “Generar una infraestructura adecuada para el desarrollo”. Bogotá. 2004.

DOMINIQUE, Lord & MIRANDA, Luis. Effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter of Poisson-gamma models for modeling motor vehicle crashes: A Bayesian perspective. *Safety Science* 46. 2008. 751–770p.

DRAGHICI, S. Data Analysis Tools for DNA Microarrays. Chapman and Hall/CRC. 2003. Chapter 11.

ESRI. Environmental System Research Institute Geodatabase. 2000. [Consultado 30-10-2009]. Disponible en: <http://www.esri-es.com>.

FURUNO ELECTRIC CO LTDA. Service Manual Marine Radar Model 1712. 1 ed. 2000.

GONZÁLEZ, R. & WOODS, R. Tratamiento Digital de Imágenes. Addison-Wesley, 1996.

HARLEY, Mitchell., TURNER, Ian., SHORT, Andrew & RANASINGHE, Roshanka. Assessment and integration of conventional, RTK-GPS and image-derived beach survey methods for daily to decadal coastal monitoring, *ScienceDirect: Coastal Engineering journal*. 2010.

HASANZADEH, R.P.R., MOGHADDAMJOO, A.R. & SADEGHI S.H.H. Optimal signal-adaptative maximum likelihood filter for enhancement of defects in eddy current C-scan images. *Science Direct NDT & E International* 41. 2008. 371-381p.

HECKERMAN, D. Bayesian Networks for Data Mining. Kluwer Academic Publishers: Data Mining and Knowledge Discovery 1. 1997. 79–119p.

INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI (IGAC). Adopción del marco geocéntrico nacional de referencia MAGNA-SIRGAS como Datum oficial de Colombia. 2004. [Consultado 12-8-2009]. Disponible en: <http://sig.cas.gov.co/node/117>.

JAMSHIDIAN, Mortaza & JENNRICH, Robert. Standard Errors for EM Estimation. Journal of the Royal Statistical Society. Series B (Statistical Methodology), Vol. 62, No.2 2000. 257-270p.

JIMENEZ, Gustavo y POSADA, Jhonathan. Proyecto de Grado: Georreferenciación de imágenes de contornos navegables mediante el Modelado de Información de Señales Radioeléctricas. Universidad del Norte. 2007.

JINXIA Y., ZIXING C. y ZHUOHUA D.. Fuzzy likelihood estimation based map matching for mobile robot self-localization. Springer-Verlag Berlin Heidelberg. 2006. 846-855p.

JOHNSON Wesley., UTTS Jessica y PEARSON Larry. Bayesian Robust Estimation of the Mean. Royal Statistical Society: Journal of the Royal Statistical Society. Series C (Applied Statistics). Vol. 35, No. 1. 1986. 63-72p.

LOEBIS D., Sutton R., CHUDLEY J. y NAEEM, W.. Adaptative tuning of Kalman filter via fuzzy logic for an intelligent AUV navigation system. Science Direct Control Engineering Practice 12. 2004. 1531-1539p.

MARFAI, Muh Aris., ALMOHAMMAD, Hussein., DEY, Sudip., BUDI, Susanto & LORENZ, King. Coastal dynamic and shoreline mapping: multi-sources spatial data analysis in Semarang Indonesia. Springer Science: Environ Monit Assess 142. 2008. 297–308p.

MARTIN, J.P. Ingeniería Fluvial. Editorial Escuela Colombiana de Ingeniería. 1 ed. 2000. 55-58p.

MATHWORKS. The Matlab Product Help. 2008.

MENDEL, M.. Lessons in Estimation Theory for Signal Processing, Communications, and Control. Prentice Hall Signal Processing Series. 58-70p.

MIRKIN, B.. Clustering for Data Mining: A Data Recovery Approach, Chapman and Hall/CRC. 2005. Cap. 3, 6.

MOURANT, Barkat. Signal detection and estimation. 2 ed. Artech House. 2005. 346-353p.

PHILLIPS SEMICONDUCTOR. Electronic Compass Design Using KMZ51 and KMZ52. Application Note AN00022. 2000.

POSADA, Jhonathan y VELEZ, Juan. Propuesta para financiación en convocatoria interna del proyecto "Sensado Remoto de Oleajes Mediante Señales de Radar". Universidad del Norte. 2010.

POZO, R. y RIBEIRO, A. Sistema de posicionamiento global (GPS): descripción, análisis de errores, aplicaciones y futuro. 2005. Disponible en: <http://www.iai.csic.es/gpa/postscript/Pozo-Ruz00a.pdf>.

RAJASEKARAN, P., SRINATH M. & VISWANATHAN R.. Introduction to Statistical Signal Processing with Application. Thomas Kailath, Series Editor. 133-173, 286-345p.

REICHERT, K., HESSNER, K., NIETO J.C. y DITTMER J. WAMOS II: A radar based wave and current monitoring system. ISOPE: Proceedings, Vol. 3. 1999.

REUTER, F. Georreferenciación de Imágenes. [Consultado 19-5-2010]. Disponible en: <http://fcf.unse.edu.ar/pdf/lpr/p10.PDF>.

RUIZ, Andrés y VILLAR, Darwin. Proyecto de Grado: Sistema de Antenas para Transmisión y Recepción, con Control de Giro Mecánico para Radar de Corto Alcance. Universidad del Norte. 2007. Sección 8.1.

SANYO ELECTRIC CO. LC79401D Dot-Matrix LCD Driver Datasheet. 1995.

SAN MARTIN, C., TORRES, S. y PEZOA, J.. Statistical recursive filtering for offset nonuniformity estimation in infrared focal plane array sensors, Science Direct Infrared Physics & Technology. 2008. 564-571p.

SERAFINO, Fr., LUGNI, C, NIETO J. C., ZAMPARELLI V., SOLDOVIERI Fr.. Bathymetry Determination via X-Band Radar Data: A New Strategy and Numerical Results, Sensors - Open Access Journal. 2010.

STEWART, J. Cálculo Conceptos y contextos. International Thomson Editores. 1999. 704-738p.

TOPCON CO. Brochure of GPS HiperLite Plus. 2010. Disponible en: http://www.topconpositioning.com/uploads/tx_ttopconproducts/HiPerLitePlus_Broch_REV_C_05.pdf.

VASEGHI, S.V.. Advanced Digital Signal Processing and Noise Reduction. Jhon Wiley & Sons, Ltd. 3 ed. 2000. 93-135p.

VÉLEZ, Juan y POSADA, Jhonathan. Informe Preliminar Final Proyecto “Actualización del SNS II”. CORMAGDALENA . Barranquilla. 2010.

YUANG, X., FU, J., SUN, H., TOTH, C. The application of GPS precise point positioning technology in aerial triangulation. Science direct, ISPRS Journal of Photogrammetry and Remote Sensing N° 64. 2009.

ZENG, D. y Lin, D. Maximum likelihood estimation in semiparametric regression models with censored data. Statist. Soc. B 69. 2007, Part 4, 507–564p

ANEXOS

1. ALGORITMO DE GEORREFERENCIACIÓN INVERSA PARA UBICAR COORDENADAS RUIDOSAS DE UN MISMO PUNTO

Sea $x_p(i, j)$, $y_p(i, j)$, $z_p(i, j)$ las componentes rectangulares de una coordenada obtenida mediante el algoritmo de georreferenciación, y $x_{GPS}(n)$, $y_{GPS}(n)$, $z_{GPS}(n)$ $n = 1, 2, \dots, n$ las sucesivas componentes de la secuencia de coordenadas del recorrido realizado.

- Para un punto dado $X(i, j)$, $Y(i, j)$, $Z(i, j)$ se selecciona aquellas coordenadas de la ruta GPS que cumplan el criterio:

$$\sqrt{(X(i, j) - X_{GPS}(n))^2 + (Y(i, j) - Y_{GPS}(n))^2 + (Z(i, j) - Z_{GPS}(n))^2} \leq ALCANCE_{Radar} \quad (5.9)$$

- Para cada coordenada GPS elegida se calcula la posible distancia $\tilde{r}(n)$ y ángulo de avistamiento $\tilde{\alpha}(n)$ al cual pudo ser detectado el punto, mediante el despeje de la ecuación 5.6 y la conversión de coordenadas rectangulares a polares mostrado en la ecuación 5.10:

$$\begin{bmatrix} d_{XN}(n) \\ d_{YN}(n) \\ d_{ZN}(n) \end{bmatrix} = R^{-1} \begin{bmatrix} X(i, j) - X_{GPS}(n) \\ Y(i, j) - Y_{GPS}(n) \\ Z(i, j) - Z_{GPS}(n) \end{bmatrix} \quad (5.10)$$

$$\tilde{r}(n) = \sqrt{d_{XN}^2(n) + d_{YN}^2(n) + d_{ZN}^2(n)} \quad \tilde{\alpha}(n) = \tan^{-1} \left(\frac{d_{XN}(n)}{d_{YN}(n)} \right) \quad (5.11)$$

Donde $\tilde{r}(n)$ y $\tilde{\alpha}(n)$ son también variables afectadas por la distintas nubes gaussianas (ver figura 27) cercanas al punto de interés, las cuales causan que la media del proceso sea vea afectada y la estimación de estas variables se encuentre desviada en cierta cantidad de metros y grados respectivamente (ver sección 6.3

- Luego se explora la matriz de distancias del radar y se seleccionan los índices que cumplan el criterio:

$$i(k), j(k) \in \left[\left[D(i(k), j(k)) - \frac{\sigma}{2} \leq \tilde{r}(k) \leq D(i(k), j(k)) + \frac{\sigma}{2} \right] \wedge \left[\alpha(i(k)) - \frac{\varepsilon}{2} \leq \tilde{\alpha}(k) \leq \alpha(i(k)) + \frac{\varepsilon}{2} \right] \right]$$

- Debido a que las matrices de latitud φ y longitud λ tienen la misma organización y tamaño de la matriz de distancias D , los índices calculados $i(k), j(k)$ sirven para seleccionar aquellas coordenadas que representan al mismo punto estimado por la coordenada $X(i, j), Y(i, j), Z(i, j)$, describiendo un proceso estocástico alrededor de la coordenada real en donde un buen estimador de ésta es la media de las coordenadas de los puntos cercanos. como se muestra a continuación :

$$\hat{\lambda} = \frac{1}{k} \sum_1^k \lambda(i(k), j(k)) \quad (5.12)$$

$$\hat{\varphi} = \frac{1}{k} \sum_1^k \varphi(i(k), j(k)) \quad (5.13)$$

- Finalmente el anterior procedimiento se repite para todas las coordenadas de las matrices de latitud y longitud, exceptuando las coordenadas que se fueron promediando en el proceso.

2. LISTA DE SIMBOLOS UTILIZADOS

R : Distancia a la que un objeto o punto de la orilla es detectado por el radar.

α : Ángulo a la que un objeto o punto de la orilla es detectado por el radar, medido en sentido horario a partir de la proa de la embarcación.

n : Total de muestras GPS tomadas de la trayectoria de la embarcación. Además corresponde a la cantidad de tomas o matrices de información de radar recolectadas.

$x_{GPS}(n), y_{GPS}(n), z_{GPS}(n)$: Coordenadas cartesianas rectangulares que representan las posiciones GPS de la ruta de la embarcación.

T_B : Vector de desplazamiento sobre una imagen cuya dirección representa la dirección del norte geográfico.

T_R : Vector de desplazamiento de la embarcación sobre una imagen, medido entre dos coordenadas GPS consecutivas.

θ : Azimut de la embarcación medido en sentido horario desde el norte geográfico y la proa de la embarcación.

d_{XN}, d_{YN}, d_{ZN} : Coordenadas cartesianas tridimensionales referidas a la posición de la embarcación, con eje Y apuntando al norte geográfico y eje Z apuntando en dirección vertical.

X, Y, Z : Coordenadas tridimensionales geocéntricas de un objeto, punto en una orilla, etc.

η : Normal al elipsoide o gran normal en un punto dado.

λ, φ, h : Coordenadas elipsoidales de la embarcación, un objeto, punto en una orilla, etc.

K : Total de coordenadas que se generan en torno a un mismo punto por las sucesivas detecciones que se hicieron de éste por parte del radar o por el agrupamiento kmeans.

$\hat{\lambda}, \tilde{\varphi}$: Estimación de coordenada geográfica de un punto mediante la aplicación de un estimador estadístico.